Download your recording (external link)

SPEAKER_04: any questions or concerns, I guess we can jump right into the presentations.

SPEAKER_09: OK. So I'm going to present first. And this presentation has a focus of using a top-level view. And basically, we were given a set of questions from Ali about things like, tell us about the project direction, Tell us about the roadblocks that you foresee for this project. Tell us about how you're going to overcome them or what the impact is of these roadblocks. And along with, give us a road map, detailed road map. So I'm going to do the top level view, and Laramie's going to do the detailed road map. So he's the detail man, I'm just the, you know, the guide that looks to see everything that is running the way it should be running. So I call them the existential questions. And so we're going to address those existential questions here. So the roadmap guidelines that we use, obviously, we have a set of guidelines that guide us. And that is to always improve security, stability, and performance. Those are the three things that drive us, right? And what drives all those three things are what the experiments need. So it's sort of a feedback loop. And like I said, the details will be in the next presentation. So just one sideline here, because everybody keeps saying, what is your roadmap? What is your roadmap? And the answer is the roadmap is captured in our web page. And here are the instructions. I'm not going to repeat them. you can read them exactly how to get to the roadmap. Because in fact, there you get the evolution, which is the roadmap, what's our long-term plan, what we're going to do in the near term. And so you can actually see what is going to be on the horizon. What I must admit that that roadmap is so detailed that people have a hard time capturing the essence of it. I will agree with that. And we're trying to very much find different ways of expressing that roadmap so that it's more digestible. And Glaramie is going to present one version of that, that perhaps you'll like much better, because it'll capture things in a much smaller amount of space. But I do want to point out that Part of this roadmap is driven by the fact that we're a very small project, actually. But we have a huge user base. And I mean, all the LHC experiments use some part of this project software. But also, we have to be cognizant that we have a huge user base outside LHC. And those people also are using the software and in critical places. And since the community needs experimental, plus all the other people outside are very dynamic, they're going to occupy our time quite a bit. And so we will be making incremental progress, but just remember that every time we do something along these lines, somebody complains about it. So progress here is going to be somewhat slow. So let's go to the foreseeable roadblocks. The one that most immediately is on the horizon is, of course, training a new release manager. This is something that unfortunately happens every five years or so. It's incredibly disruptive when it does happen. And so that is a definite roadblock that we're seeing that is going to cause some amount of disruption. The other one is, of course, roadblock is user base is growing. But the available effort that we have is very variable. We really never know exactly how many FTEs we're going to have to work on anything. And even when we start work on a particular project, we're not guaranteed those FTEs are going to last there long enough to complete it. So this is a problem with a community-driven project. I mean, the labor is, shall we say, unpredictable. The other thing that we just experienced in the last year is the onslaught of AI discovery and AI code

submission. So when I say AI discovery, what I mean is people launching Cloud or whatever, looking for security issues. Now, that's good. I mean, we want to correct security problems, but people have to remember that if we all of a sudden get a report of a security issue, all work stops until that security issue is addressed. Because we don't want any kind of exposure that's going to basically compromise data. Now, in terms of AI submissions, as AI is supposed to actually alleviate human labor, it does not. because we're getting new AI submissions for people that didn't actually submit things to our project. And now we have to do really good code reviews and more scrutiny because these submissions tend to be far more complicated because the AI can write circles around anything. And so we're trying to grapple with that problem. One of the ways of grappling with that problem is having AI review the AI. And that sometimes works and sometimes doesn't. The other last thing is that project continuity is not assured. We keep bringing this up every year. And I want to point out again, we're the only project of this size that has no full-time staff. And I don't know how you actually can guarantee continuity without full-time staff. So what's the bottom line? Impact. The project has always faced these roadblocks, actually, and it's amazing that it's held together for the last quarter century. And I can only say that's a testament to the community dedication to this project and the people who work on it with their zeal and knowledge. Without those people, we'd be dead out of the water. And so kudos for all the people in the trenches. So the biggest impact we see from these roadblocks is significant slips in planned roadmaps. We can do a roadmap. It looks great. But something happens, and one of these roadblocks just causes us a slip. Now, this is the worst case scenario. I can always point out that there's always mitigations we can employ when we get to a roadblock. But you have to understand that only reduces the impact. It doesn't necessarily eliminate the effect of that rule of law. So now we can have a discussion if you'd like, or we can go to the second presentation.

SPEAKER_06: Thank you very much, Andy. I would suggest that maybe we, if there are no immediate questions, Um, on this topic, I would suggest we go on the next presentation and we take the discussion. Okay.

SPEAKER_09: Next presentation.

SPEAKER_06: And maybe at the end of this, please remind to mute yourself. Otherwise it's complicated. Thank you.

SPEAKER_08: It's the first one. No, no, no, no, no, no, not this one, Andy on your. Part? No, that's Panos. You can leave that loaded for later. No, I can't because I just destroyed the screen. Go back. It's on yours? It's the first link.

SPEAKER_11: Yeah. First link. Yeah. Sorry. OK. Let me get out of the way. I'll give you the microphone. Thank you. Okay.

SPEAKER_08: Thank you, Andy, for the presentation. I made a set of slides here. I'm not going to go over all the details. A lot of it was covered by Andy at the beginning of the workshop, but I wanted to leave as a reference. This is just as like to say, X3D is a very important project for not only the high energy physics community, but also beyond. So since Andy already made this,

already talked about this, I'm going to skip. So this is the previous, what happened before. So Andy mentioned we had some, doing both of them while people are still using 5, they contain the fixes. So you should be down here in 597 and 611 if you want to be secure. And of course, we didn't manage to do 598 and 620 before the workshop. But there are still some things that we need to finish. But they're going to come out soon. A lot went on in these releases. There are fixes for this big test with 40,000 transfers in FTS, in addition to the security fixes. So really, if you have the chance, upgrade to the latest ones. And it brings our focus. So as Andy said, security, stability, and performance is the top priority of the community for XRoute and is also our top priority. So this is kind of the baseline of what we do. So here are reminders of what happened in 6.0, 6.1. I'll just say the relevance of the items. So Andy mentioned that we had server-side support for file cloning. So the impact of this is that this allows Eos to implement read and write support for erasure-coded files. The importance of replacing Davix with Curl is that Davix and GFile, they're in this decommissioning phase. We want to move to more modern. stack, concentrate things in xroot, and allow the client of xroot to be the client of the grid, so that's the importance of that. The state operations with tokens, this is part of the migration to token theme. Data integrity, I guess I don't need to explain, so you need to ensure that the files are there with the right checksum at the end. and I skipped 6.1, but it's very important to have streaming real-time checksums. Mihaj allows testing. OpenSSL support is important for the future. But 6.2, so we're going to have HTTP support in the client. There's a talk tomorrow about it. Another talk tomorrow about it on the TAPE REST API implementation within XRDFS. So this will be coming very soon. It's already merged into master. So this will allow you to basically use XRDFS like you use GFile today, but with a similar syntax in the command line, with the HTTP URL, but just with our client. And this goes on the theme of consolidating things into the XRD client. OK, so in the next releases, There will be monitoring, so you had long presentations from me about what we're doing to support metrics, logs, all the monitoring features that are needed natively within X3D. There will be also support, native support for OAuth2 tokens, which allows to connect with SSO tokens to XRoot. And we'll continue the evolution of the Python bindings to support this migration to GFile. And 6.4 and 6.5, we also have items planned. But let's focus on the formats that WLCG wanted, which is Key Mile they have. So for 6.2, we have the client, the concentrations of the features in the XRD client, which enables to phase out GFile and Davix and make XRD the client in the future. So the streaming time checksum is to avoid this re-reading the file after writing. The storages will free up a lot of activity on the disks, on the storage endpoints. So these two are done, they will be released very soon. And then coming next will be the monitoring. So the solution that we came up with, and this will be really plug and play with the monitoring stacks, like Cognitive Monitoring stacks. And also here, the use of SSO tokens, the use case in mind that we have is SWAN, where For improved security, they want to replace their Kerberos setup with tokens directly from Keycloak, and everybody else will be able to use this with not only Keycloak, but also with GitHub tokens, OAuth2 authentication. For the future, like the longer future, we want to match the support of KPress API also in the server, so there is already some let's say, simple implementation, because we needed this for testing that the client actually works. And we want to evolve this into a full implementation of the Takebacks API

with the Prepare plugin of XRoute. And this will enable vanilla XRoute integration with Takeback cams. And we plan to have this around June next year. Then the big item that is forward-looking towards Halloumi is the IOU support, which will allow asynchronous operations on the server to the disk, so that you don't have to have the thread holding the synchronous operation and then returning. So you have to have many threads if you want to have many operations in flight. You can have just a few threads, and they can manage many, many operations in flight. So this probably will go into X3D7. which will be in a couple of years from now. And the other two big things that we look into is XRDHTP is very good, serves us very well, but it's only HTTP 1.1. And there are certain things that can be improved with moving to HTTP 2. We don't know, HTTP 3 doesn't buy much in terms of features support beyond what HTTP 2 already provides. But it does allow to improve performance of multi-string transfers and have more synchronous operations. So that's also something that we plan to implement. So replace the kind of custom implementation with a library like ng-http. And that should also come together with Xubuntu 7. And even further down the line, we think maybe if we want to evolve both EOS and XRoute together as a parallel NFS solution, then we can replace the fuse mounts with NFS mounts, and these NFS mounts can use the kernel clients, so that's something that is very complex now, it's kind of buggy that we can just use something very well maintained from the kernel. But I think the key message here is that even if we have these big, big items, XRoot today is pretty much ready for run 4. You will use more CPU if you don't have this asynchronous I-O with I-O U-ring, but it will work. We do have enough room in our storage nodes, the tier 0, that even with the increased data load, like the load on the network, the work we have done before for 100 gigabit optimization is enough I believe that to, I mean, the server was already ready. The client needed the bug fix, but it should be ready for run four now. And then we have this kind of, this is a slide I took from the OTF on March that started as a joke, but became a thing. that we're really going to become a cloud-native thing. So I already improved the system D units to make it more secure, to make it easier to configure with the standard parts for HTTP. We're doing the monitoring now. We're going to do authentication next. So this is a thing that rules, you know, better integration with Kubernetes. So you can see, you can expect us to provide like an operator or help charts in the future for XRD. And, and that's it. So XRD is important for the community and we listen to everybody and we'll do our best to improve the project to satisfy everybody's needs.

SPEAKER_04: Okay, thank you very much. We can open the floor to questions now at this point. Andrew Mello.

SPEAKER_08: If you want to hold the questions after the panels, we can, but I think we can have a couple of questions.

SPEAKER_05: No, no, we want to have the questions now and then we will have more links for the monitoring.

SPEAKER_01: Okay. Okay, hello. So just briefly, you mentioned... There was this mention of token support, right? And you mentioned, you know, CERN tokens and GitHub tokens. Is this a regular, you know, OIDC endpoint? So you could point to both Keycloak, GitHub, or, you know, some other solution, or is it specifically those that are supported?

SPEAKER_08: We want to support standard OpenID Connect. So then any OpenID Connect solution should work. So this is something Andreas is actually going to talk about it tomorrow as well. So if you're interested, tomorrow morning. So let me block.

SPEAKER_09: No, no.

SPEAKER_11: I can use the block. No, no. Blue square. Oh, really? Why did you have that? OK.

SPEAKER_09: So Andreas is doing that work for OpenID Connect. It's very much a research project right now. We're on the third iteration, I think. And he'll present what he's come up with so far. I think he's made great progress. It's not there yet, but it's pretty close. I think it'll easily make the deadline here if we have a decision for it.

SPEAKER_08: Yeah, yeah, December. So it should come together with the monitoring.

SPEAKER_01: OK, thank you very much.

SPEAKER_06: Can you get the mic, please? You are a bit far. Difficult to understand.

SPEAKER_08: Is it muted, the mic?

SPEAKER_11: OK, I'll speak louder. It's OK. OK.

SPEAKER_09: Oh, yes. Yes, we are getting some funding from FCC. Again, I wouldn't call that assuring. It's very small.

SPEAKER_06: What did you say, Andy?

SPEAKER_09: It's not something that you can count on per se. It will take us over for a couple of years, but every year there's a decision point on whether that's going to continue.

SPEAKER_05: Okay, to stay on the same topic. Okay, someone from the room, please go ahead.

SPEAKER_06: Maybe the mic in the room is not on. This might be because I hear very well, Andy. I hear the others very low, but I heard the... I heard Cathy's comment.

SPEAKER_08: I'm bringing my speaker microphone to people.

SPEAKER_06: Or maybe switch on the other one.

SPEAKER_08: Can you hear me?

SPEAKER_06: Yes.

SPEAKER_08: So is this better?

SPEAKER_06: Yes, it's much better. Yeah. But we heard Cathy. It's good.

SPEAKER_08: I think it was muted. Okay. I think it was muted before. Okay.

SPEAKER_06: I have a question related. I see that you are, you have releases basically three months. How easy and how fast is the deployment of the, um, of these releases?

SPEAKER_08: Um, so this is actually something that we changed in the last workshop. Um, we used to have releases every six months. And that created a bit of friction with some people if they didn't have a certain feature ready on time and so on. So I decided to make releases every three months instead so that with a cadence of three months, if somebody wants to have a feature

in that doesn't make it, it doesn't have to wait so long to get it in. And the deployment, I cannot speak for other sites, but normally with Within CERN, we picked right away the new releases. So there is a version of EOS that is already using x3d6. And we offered our repositories, our official repositories with RPMs and Debian packages for major distribution. So the latest Debian and Ubuntu, the latest Fedora, the latest Alma. And if you use those repos basically on the same day of release, you can upgrade. or you can try the release and so on. And we have the stable channel and the testing channel. So my plan is when I release 6.2, likely next week, maybe it will go a bit over to the other week. Then right away, people can download and install the package. It's really simple. The instructions are on our website. So usually, if you don't need to make use of a new feature, it should be a five-minute operation.

SPEAKER_06: So the specific question is the last line on this open idea of integration. Basically, do we need to ask sites to migrate to it in any moment, or we do it when the sites want to do it?

SPEAKER_08: No, this is something that decides. So this will be a new plugin, so it will be there, available. And then they will have to add a couple of lines in their configuration to enable it. And we'll provide documentation explaining how to enable it when the release is out. So you can upgrade to the release. It will be there. You don't need to use it. But if you want to use SSO, for example, the people running SWAN, they will want us at certain tier 0 to deploy this right away. And most likely, when we upgrade EOS, that will be part of, like, when we upgrade to EOS that contains X3D 6.3, at the same time, we will enable this configuration for the plugin.

SPEAKER_06: Thanks. James, I have another technical question. and which I would like to ask Guilherme and Andy and the community. You mentioned that, I think the most important thing that you mentioned is that you have a solid plan. It seems very interesting. And also you mentioned that the reality is that even today, you believe that xRudy, it's basically ready for iLumi, which is very good. And there are still room for improvements. It's also very important. The question I have is, what is the metric you are using to tell us basically that you feel that you are ready for RAM4. Is throughput, is what?

SPEAKER_08: So if you look at the slides of Gustavo, he showed that the local host, a single server can reach well beyond 100 gigabit per second serving from an SSD. So our storage nodes are far from that right now. So the hardware, what we try to ensure is your bottleneck will be on the hardware you buy and not on extra deal. So this is our mentality for this. So if you buy expensive enough hardware or enough hardware of cheap hardware, then you will have the performance you need. The software is not going to be in the way. And of course, there are places where there are spots that need to be fixed. I think for Xcache, for example, Mateus is going to work on making it work better for analysis, which was something that was presented this morning. On the server side and the client side, what we will do is we have, we're working, we look at SORI to get access to 400 gigabit networking, and we were going to ensure that extra D works well there until you saturate the link. So if the hardware can do it, X-ray D will be able to do it. This is basically our plan for HLUMI. And that's why the IOU ring, in that case, it might be necessary for 100 people.

SPEAKER_09: So let me add one more thing, because that's our guidelines and performance markers for performance, right? How we do in performance. Remember the stability and security, those are the other two. And we have also markers of whether we're doing well or not, or how we gauge that. In terms of security, the gauge is how quickly we can get a security fix in through the CDE process, if we need a CDE process. And generally, we do fairly well there. In terms of stability, the way we do the markers, or how we measure that, is how quickly we can reduce the number of open tickets. And that's a very visible marker. And I'd say we're not doing as well as we would like in that area. So that's where it is.

SPEAKER_11: Thank you.

SPEAKER_08: Fair points. We do have over 100 tickets open now. We've managed in the past to bring it down to like 80. But yes, as people filed more bugs and so on, I hope that, I mean, I try to focus on bug fixing most of the time, but yeah.

SPEAKER_09: It's just a matter of having the right people at the right place at the right time to address those bugs. Yeah.

SPEAKER_08: Everybody loves developing new features, not so much fixing bugs.

None: Right.

SPEAKER_08: All right, maybe we can let Panos talk about the monitoring.

SPEAKER_04: Yeah, let's do that.

SPEAKER_08: I had two very long talks. I think Panos did an excellent summary of it. And we're going to work together to get the solution working by DC27.

SPEAKER_11: Thanks, Guillermo.

SPEAKER_03: So is this being shared, and you can see and hear me? I hope so.

SPEAKER_05: Yes.

SPEAKER_03: OK, thanks. Go ahead. So I'm back with our favorite topic, which is XUD monitoring. As Guillermo said, we had a big session yesterday with four talks on XUD monitoring. So I'll try to recap a bit, also for the OTF audience to know what we talked about, and then lay down the plan on how we get to PC27. to what extent we're going to have it should be monitored at this point, and then how we move forward after this. So what we covered yesterday was two complementary solutions, I might say. So we have the OSD collector, which was developed by OSD, but then improved by the WSG operations and the monitoring. So we got this Go collector from OSD, improved it and validated it at CERN. We put it in production at CERN for all LHC experiments. And now we want to move forward with deploying this at many different sites for DC27. In parallel, Guilherme presented a native XOD monitoring solution that XOD team is developing. This consists of two different parts. The one is the matrix part, which is a Prometheus matrix built into the server. Then it's a monitoring collector part, which is essentially the same thing. per file, per operation recording. And then it also includes error reporting, which was not there until very recently. This is still in a development slash pre-production state, and it's planned to be released in 6.3, it should be released in the end of the year. So both of those approaches

have the same goal, have bio-agnostic trustworthy XoD cancer monitoring. focusing a bit on the OSD collector. So what we did is we validated it and we're running it at 7. There were a lot of improvements that were done to it. First of all, on the observability front, we have metrics on how many packets we receive, how this is adding to records, how the collectors are performing, and things like that. The configurability, we had to make the collector configurable enough so it's easy to deploy to different sites. So we have site-level configurations now, which were not there before. In terms of enrichment, we have to bring things in like the activity which was missing, the authentication method, the separation between traditional and vector reads. This is all there. And we also did a lot of work on the port topology resolution. So what you get from XUD records is IP addresses, but what you need is a site name. So we have to do the resolution. We do this via click. We can also differentiate between LAN and WAN traffic and also by actually internal traffic, I mean real traffic that's coming from the storage system itself, and it's not initiated because of some user activity. And also some bug fixes, of course. We have five collectors running at CERN, four for the latency experiments, plus the CMS pilot one. And we get 120 million transfer records per day, so not bad. The way to validate this was to compare it against Alice monitoring, which is the only reliable source that we had. and there is a mean difference of 3.5%, so we are very, very close. Also, we fully understand why this 3.5% is there, so we deem this validated and ready to go. So we have a golden release available, and of course there are some changes that need to be made, some streams, some last-minute bug fixes and improvements, but this is more or less in a stable state right now. And here is the product of it. You can see this is a 24-hour the server network. So jumping to the other solution that we talked about yesterday, the native XOD monitoring. As I said, this has two parts. Yigermi developed the XOD metrics part, which is Prometheus-like metrics, detailing how the server exactly works, exposing the health metrics and the overall accounting of the storage system. really for live metrics. For transfer accounting, there's actually a monitoring collector, which is a UDP collector that stitches together the packets and produces a record, like the other collectors. And the idea is that, yes, you can have these records, push them to open sets, and visualize them as you want. And the very nice new thing is errors monitoring and client monitoring. So finally, errors are now in the fStream, so the collector can grab them. This also would work for the SD collector. It can be everybody can do it on the fStream if the server is running 6.3, I think, onwards. And then also, what we're going to implement is client-side metrics. So this is going to be client-side monitoring for newer clients, of course. And as I said, this is in pre-production. The release plan is for the end of the year. and 6.2. There is also DCAS, of course. So it's not only XOD native. DCAS sites that run transfers or operations via their root protocol follow a completely different model. They have their own flow independent of the collector we are talking about. Today, to do this, one needs to enable the Kafka bookkeeping system. access and deploy Kafka at their site, and then there is some logic that pushes records centrally. For Fermilab and Peak, we are already getting data out of the DKAS sites. For the rest of the sites that are not running the latest DKAS version, I think we won't push to get any information from them unless they really want to, because from DKAS version 11.2 onwards, there's a new way to get this directly without the need to have Kafka at the site. To do this, we need an improved collector. There is already ongoing work at ESRI,

so the DDoS developers are helping us with that, and we hope we have something soon. And this is not only writing a new collector to get exactly the same information, but also we add new information, like the file names, which are needed for data popularity for CMS and LHCB and probably others as well. And here it's worth mentioning that we should try and we will try our best to use as much as possible a common model for this record. Doesn't matter if a record is coming from a DICAS monitoring or from a OFD collector or the XUDIN81, they should try to follow the same schema as much as possible. And next for DICAS sites is, of course, moving on with the deployment campaign, enabling this flow to all sites. So that's the deployment plan for the OFD collector. We have the golden release, which is technically out there already and we want to deploy to sites other than CERT. The first volunteer was RAL, so thanks a lot to Giotis and Katie for their volunteering. We plan to deploy the collector there fairly soon and of course brace for all the needed changes that we will discover when we deploy it there. Then, because of the next months in the autumn, we want to have a lens centrally and also get the D-cache version that does monitoring for versions 11.2 onwards. So then, towards the end of the year, we can start deploying this to as many sites as we can. I'm sure we won't be able to deploy this, the whole grid, by this 27, but the plan is to have this to at least a few major sites. And then by DC27, we would very much like to have actually the graphic monitor at as many sites as possible, including all the USPR2s that are already running the OSG collector. So we get some information out of them. There are some things missing there, like a proper topology resolution for WSG topology, of course. But the data are there. So in terms of throughput, we can plant them on top of the rest and have this recorded. On the native collector side, the XOD team will continue developing this and improving it. And the very good thing is that we can run those two in parallel. So you can have any XOD server sending data to two different places, sending streams to different collectors. So we can have the OSG and the native collector running in parallel. We already have this in and then for DC27 we would like to have those running in parallel for at least seven years. So in terms of long-term view, the obvious thing, the elephant in the room is that we have two collectors, right, but they do more or less the same thing. So the native XUD collector I think is technically the best solution if the design is looks very, very good, and the prototype looks solid. So we get exactly what we need out of it, and you get the benefit of having a native collector that is shipped together with XUD. And also, as XUD evolves, the collector evolves together with it. So why we didn't do this already? The thing is, we needed to have something for DC27. So when we started working on the 3D collector, it was the only available choice that we had. So for DC27, we'll go ahead with it. and we'll continue to develop it and try to make it fit our needs. But then after DC27, we can revisit this topic. Of course, first of all, we need to have the XUD native collector being finalized, merged, validated, and extended with functionality that WSG needs. Things like topology solution are not yet there, but I'm sure it will be easy to add. And then, of course, if we want to switch to it, that this will be supported. So it's great to see that the monitoring is finally becoming a first-class citizen in the X2D world. But as I said, now, if you run monitoring, you need to update it as the software goes on. So in parallel, we continue the excellent collaboration that we have. We evaluate both connectors and we try to improve them both. And I think that's all. This is my summary. So the X2D Transfer Monitoring at CERN

is considered reliable running the OSD connector. The plan is to have data challenge done with the OFD collectors to as many sites as possible. We start with RAL, deploying it there very soon, hopefully, and then we get to more volunteer sites. For DIGAS, we're also progressing. We rely on the DIGAS developers, of course, to provide the solution that we need. But they're already working on it, and I hope this is not going to be a lot of work for them. And in parallel, we continue working on the native collector and try to make validated and production ready. And I think that's all I had. Thank you very much.

SPEAKER_04: Thank you very much, Panos. That was very interesting. Should we ask for questions?

SPEAKER_03: Yeah. Katie already has one from the room, so I don't know if we- Please go ahead. Does the cube work after all? I don't know.

SPEAKER_11: We can try. We can try.

SPEAKER_10: Correct me if I'm wrong.

SPEAKER_03: No, it doesn't work.

SPEAKER_10: It doesn't work? Yeah. OK. So the question is for the xRuby developer. From my memory of the past, xRuby would not like to feed more than, say, two monitoring streams. Have I understood this correctly and have this changed?

SPEAKER_11: So yes, so the current support is two different endpoints. Alice has requested to have a third. So I discussed with Andy, signed off on it. We will probably add a third. We shouldn't add too many, because it's taxing on the server. So you can hurt performance of the server if you enable tracing and send too many places.

SPEAKER_10: Yeah, three would be good if we're having these two plus sites like to have their own.

SPEAKER_11: Yes, we can have a forward mode in the collector. My collector has a forward mode, so that's in my mind already. You can send to one collector, and then that collector gets into all places. James wants to, remotely wants to speak. James?

SPEAKER_04: Hi there, yeah. Thank you very much for the presentation. If you could go back to that timeline slide. I forget what it was called. It was slide seven or eight. Yeah, this one. So yeah, this looks like a very good timeline through DC27. And of course, DC27 is very important. And I'm really pleased to see that we'll have at least some monitoring of the X3D transfers during this challenge. I'd encourage you to, understand where deploying the monitoring would give the most completeness to the challenge data. I mean, if we can target maybe large sites like CERN or Fermilab or BNL, for example, that might be good. My question is a little bit more long-term. So I understand now, okay, you have these two solutions. You have the OSG collector and the native solution. And at some point in the far future, or at some point in the future, I'm not sure when, you have what we would call a decision point where you're going to, you say that you favor one over the other. It'd be good to know when you were going to make that decision. And it would also be good to understand at what point in time, you know, what year do you think that in the

WLCG infrastructure, we could have more or less a complete understanding of the XRouteD monitoring? So, I mean, one of the challenges of the users, the experiments effectively, is really having a global picture of their utilization, of their file access, for example, And it would be good to try to understand and really write down, when do you think we could achieve that? And what are the challenges or the decisions? OK, you have these two solutions. So maybe there's a decision that needs to be made there, maybe not. But it would be nice to understand that sort of end game of having fairly complete monitoring across WLCG. When do you think you can achieve that? And what do you think the roadblocks are?

SPEAKER_03: for the first part for the decision, putting this after the data challenge makes sense because we want to reach the data challenge with as much coverage as possible. So we focus on the solution that works for us, which is the one for now. And then by then, if the exodinative collector has been released in the end of the year, then by March, it will have two or three months time to run. So we would, at this point, know how well those two run. Also, data challenge will help us compare them, essentially. and then we can move forward. Now, on how we get these two sites, once we decide on the collector, we need to deploy this everywhere, right? And this is always a hard part. I need to think of the details, but unfortunately, I think the answer is always that the first 80% of the sites will be very fast and easy to deploy, and then you have a long tail that we need to run behind. But as you said, we start with

SPEAKER_09: Yeah, I want to provide some more context into this. So the OSG collector was problematic for us, because in order to package it with Xroot, we'd have to package all the Go dependencies. And we found that unacceptable. We presented that to OSG and asked them, you understand our position here. And they said, yes, we understand your position. We would not recommend you packaging GOME dependencies. So they're well aware that there's a competing collector that will be actually distributed with XRoute. Now, the question comes in that it's very easy for European sites to just go with the native collector. I don't know exactly what the transition will be in the US sites. So I can't vouch for that, because that's really dependent on what OSG wants to do. But for the European sites, it's a fairly straightforward path.

SPEAKER_08: Yeah, so just to give the point of view of release manager here, my goal is to put something out in 6.3 that is really production ready. And between now and then, so we already have it pre-deployed. I have your CMS monitoring going on. So I'm using this time to refine it so that when it's available, everybody will be able to just upgrade X3D, start a new system D unit with a simple configuration and have monitoring.

SPEAKER_06: This monitoring, what does this mean? Does it mean that in the, let me say the end goal, the WCG dashboard, What's the steps to reach the, to have the files monitored, the files touched, accessed in each file monitored up to WCG?

SPEAKER_08: So, of course, that as part of my work, I had to deploy the full monitoring stack. And I also wrote a deployment guide. So if you don't have any monitoring, Then you can follow the deployment guide, install a few VMs, and have monitoring working fairly quickly. But for everybody that already has some monitoring, that's, for example, if you have an open telemetry

collector, or if you have OpenSearch, it will be really plug and play. So you just point the collector at the OpenSearch instance with the channel that you created for X3D. And then the dashboard for Grafana will be also distributed as JSON to be imported into Grafana. And this is what I'm working on refining at the moment, like these dashboards, so they show all the proper information and so on. So I want to make it really, really painless for people to really configure a lot of files. plug to their existing monitoring stack and have it work.

SPEAKER_06: So my question is a bit lower level, most probably. But my question is, if today we have an almost reliable monitoring provided by the OSG collector, as shown in the presentation earlier on and today. And this one is basically done through the four or five collectors that are running at CERN. Then whenever one of these sites is already monitored, we will upgrade to a release that will contain the monitoring from native monitoring through the what will we get the double counting what will it happen i see by the way that there is a and there is moria with i don't know if you want to answer i understand i understand your question so right now we run the osg monitoring and my monitoring at the same place but you have to choose where the data comes from so there is no double count

SPEAKER_00: Okay.

SPEAKER_08: So I have my dashboards and Panos has these dashboards. They're based on the same kind of monitoring streams that are coming out of the server, but they're in different places. So in the other places, unless you, so you will configure the two destinations for your monitoring. When you enable the new one, it can run alongside the old one. And then eventually, you choose to keep the one you want. So I think it can be very gradually deployed, as long as you don't need the third. I hope that I can put the bug fixes, error monitoring into 6.2 so that we can use even the pre-production collector with production extrude and have the monitoring of the errors. and have the OSG collector do the error monitoring between, so that we have that information, even for EC27 already.

SPEAKER_06: Thanks. Borja, his hand raised. Borja.

SPEAKER_07: Yeah, I just wanted to add that there is something that has not been discussed and we need to take into consideration. With the new XRudy native collector, it follows industry standards with the output to OpenTelemetry.opensearch, which is, in my particular point of view, very correct. But the end goal for this data for WLCG is to show it together with FTS data in order to have the full picture of the transfers across the WLCG. Right now, the OEC collector, it already outputs the data in a format that is fully compatible with FTS. So merging this data together is fairly easy. For the native collector, there is going to be an integration effort that will need to be done, which is something Again, we should do because I think the way that the extra the native collector is producing monetary is the correct way. It's some kind of isolated product and you can run this outside of WCG. But to show this together with FTSA and WCG, we need to work over it.

SPEAKER_08: Yes, so my plan, I did the, so I already have open telemetry and open search as separate things if needed. I can add the third, like WLCT output format. It's exactly the same as you can.

SPEAKER_06: Thanks. I think this discussion is super interesting. Maybe what I would suggest is that a few of these details are summarized in three or four points, and maybe a milestone or a tentative milestone can come up, because this one seems to be very clear. A decision point, as James was suggesting, can be I don't know, three months after DC 27. And then there might be maybe again, other nine months. I mean, we can take time. It's not a matter of being in a hurry, but I think it's important to write down that we have a plan and where we are going to. This is my point. I see Andrew also has his hand raised.

SPEAKER_04: Yeah. And just to give you some time pressure, we do need to finish writing and submitting the WLCG technical roadmap by the end of the year. So it would be good to try to write down this plan in the next month or so. But yeah, let's go to Andrew.

SPEAKER_01: Hello, I'm curious about what are the hardware requirements of the collector itself? So, you know, is this, you know, you could run it on a small VM or do you need something more substantial? You know, let's say for like a large-ish site, what does that look like?

SPEAKER_03: I should have kept the backup slides from yesterday's presentation. So for the OSG one, I can answer, then Guillermo can speak about the native one. So what we realized, so as I said, we have four deployments, one for each. And then for the Alice collector, we use around four gigs of RAM and then the storage is not so much.

SPEAKER_04: I'm sorry.

SPEAKER_03: It was, how many? Uh, four or five gigs of RAM.

SPEAKER_01: Okay.

SPEAKER_03: Alice traffic. So it is not a big site with one collector. Maybe it will need like 10 gigs of RAM. I think that's my impression.

SPEAKER_01: Okay, that answers my question.

SPEAKER_03: If you run one collector for all your traffic, you can always split your items.

SPEAKER_08: Okay, for the native collector right now, I'm running it on a pretty small machine with six gigs of RAM and I'm setting, I have a memory limit where it would just delete the very oldest data if it needs, if it goes into memory pressure. So with, I'm receiving information from about 270 servers now. with tracing enabled, and it's just going fine. So I think that the memory requirements is not that big. But later on, when this comes with xroot, and it's just a new systemd unit that can be run together with the server, I think that can be run distributed and not have a like, we only have the central endpoint where all the information is sent, and then the collector can run together with the server. And the decision is easy to either send via TCP with a shoveler or just start the collector mode. Once it's together with xroot, there is no extra step, right? You just start the new xroot unit, and then you can have the monitoring. So we will not need the extra hardware. But even a small VM now is enough for USEMS, USPublic, and USPilot together.

SPEAKER_01: I see. Thank you.

SPEAKER_04: OK, great. Any other questions? All right. I guess, Ali, we could probably try to maybe try to wrap up if there are no other questions. I think we've had a good picture of the xrootd and the xrootd monitoring development plans. over the next couple of years. Something to work on, of course, is a bit of the monitoring plans post-DC27, and we'll be in touch with Panos and Guillaume about that. I don't know, any final words, Ale, or any final comments?

SPEAKER_06: I would really like to thank you very much. Very briefly, the XRudy folks, the community themselves, we heard Andy your talk very clearly, loud and clear on the problems, the challenges of the past 25 years and the personal power issue. We might need to discuss, but this is not an open technical forum to discuss, if and how we can help. We might need to have a chat with you really on trying to understand which could be the places from where we can try to to get a bit more support. But it's not for this open technical forum here. So thanks a lot for having done the effort of coming down with all these milestones for the next years. And that's all.

SPEAKER_11: Thank you.

SPEAKER_04: Thank you, everybody. Bye.

SPEAKER_11: Thank you.

SPEAKER_04: And we close on time this time.