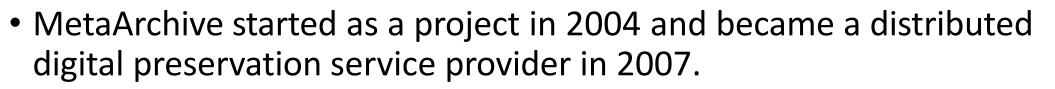
PROJECT MYCELIUM: NEXT GENERATION DISTRIBUTED DIGITAL PRESERVATION

Nathan Tallman Penn State University <u>ntt7@psu.edu</u> @ntallman@digipres.club Glenn Heinle KeeperTechnology glenn.heinle@keepertech.com

Library of Congress Designing Storage Architectures 2023

Introduction





- Since the beginning, MetaArchive has used Stanford's Lots of Copies Keep Stuff Safe (LOCKSS) to provide bit-level preservation.
- LOCKSS is great for tightly controlled networks and static journal content. But MetaArchive embraces administrative diversity with distributed network management and members deposit a wide variety of cultural heritage materials, in a variety of packaging standards.
- LOCKSS has served us well but is now inhibiting network growth.

Triple Bottom Line Sustainability and UN Sustainable Development Goals

- Triple Bottom Line Sustainability
 - Environmental Sustainability (Planet)
 - Economic Sustainability (Prosperity)
 - Labor Sustainability (People)
- UN Sustainable Development Goals
 - Goal 9: Reduced inequality within and among countries





BUILD RESILIENT INFRASTRUCTURE, PROMOTE INCLUSIVE AND SUSTAINABLE INDUSTRIALIZATION AND FOSTER INNOVATION

Research and Development

- Partnership with Keeper Technology
- Multi-phased Approach
 - 1: Requirements gathering, system design
 - 2: Prototyping, proof-of-concept
 - 3: Deployment of a pilot network





Discovery

									P			
	Inpetialist	Anna daurka pro-probrat nadarite te tepeta dalar antari teta a prostantia artist	What itsel of granularity is declear for series and district?	Which holding out had have using the second logals/Pa secon ²	Mar tod of (bac-boke) Metados O pos order 5 heap Mathie d'acado?	Poer Segar en lite active algoris poe seur to satisfi is the stronge andre search	Weak yes the new action one believe above and device the recent?	official for galling access to the president system?	Wandi defining a salasikalar Rationa kar parka?	Passe fait as a Bits 18 stand pass surger had it influences applying particular is a presentation system.	Head you to althe to perfoliate it a Minney Minney part works and Material same if you, phase promises as and phase promises as and phase promises as and	Continue Continue
100,00,00	Avera interva	design of the second se	ter seneralista en la secola en la lectra destructura i activativa en la lectra destructura en la secola de la lectra de la secola de la secola de la secola la factore de la secola de la secola de la secola la factore de la secola de la se	Paramenta and a second and a se	ngelie of an an and a state of a		for all the solution of an or phoneses, rate, and observations and the	The set of the second set of the second seco	N. Alf and an order of the second sec	The constraints are an expected with start the second starts in the second starts are expected with the second start time in the second start time start are start and start an	turing Respire Independent (Independent Jacobs (Independent) Independent (Independent) Independent (Independent)	
	unandy of Locksfie	th an off, so with south large of Solar solar links in	N ADALAN ANY MANAGEMENT		Contract in the contract of periods in the destruction matters of the second entrusts addresses the second entrusts of the termination.	All added and a set finite in the set of the point of the point of the set of the set of the set of the Point of the point of the set of the	Anna Landaritani albagini bi Manani albagini bi Manani albagini bi	-	Nale, 5 or all some also the	No han annan sanan lanak dara sana In ini kana diri cika bala sa sa sana Ini ang		
inite for a	Andreder andres (serve	No statute et a signe, sociale et a si secondario et al si de la superiori de la signe estatute estatute et al si de la superiori parte estatute estatute et a signe parte estatute estatute estatute		County county by approximation .	The "generative consistive" and of variants, respectively and properties of the poly-	The principal set of the principal set of the set of th	unit.		******	*****	~	
	linger figte University	Frysler Rody server wi		Provide National Activity of the encoded in the last solution, or the descent and the solution of the provide encoded in the last of the provide encoded and the solution of the encoded activity of of encoded activi	And a solution of the second o	CONTRACTOR OF TAXABLE	propriet and a single state of the					To all of the Version Constants to any other states of the Constant of Constants of the Constant of Constants of Constants of Constants of Constants of Constants of Constants of Constants
101.00.00	Arter standy	Reference of a second s		The ball of the second se				and an and the second s	2	en na che cola se an an haddel han Naci il fa forai	-	
	Ola (standa (desta	kolonisti konkurat Malanisti konkurat Malanisti kara				lation in a situation of a set.				In the second color of the second sec		
	ingen funt. Untersty skruter	National and charling with Role in Spinistics	Process control for the process of the pro- section of the process of the proces	NO PERSONAL SPECIAL PROPERTY.	No and Alexandra, the Alexandra of Alexandra, the architecture inclusion along the architecture and architecture and the Alexandra arc and a	Anone is an in the second seco	where the in the second second	An and a state of the ballion and	Coll for an an out-that it before the foreign of the corre- solution of the correspondence of the correspondence of the correspondence of the corresponden	i far sen nel sen felt. Badi alter for serie dels far la sent farattina finale del fasta finale dels sent Paren		
	Ness Trans Sector of a Hereite	No. 6 In case of the second se		the second secon	No. No. White an other is a state of a state of the state. It is maintain a state of the state o		n ng chi ya shahan ƙwa Malumi ƙwar	ar an a set one of the second second	Age of the second secon			
	AC Wandsal Ukana Milana Manal	No los del control militar la report a processaria de control militar de la port controlar de control y control de alterna del anticipar de la portecida de portecida de control de la portecida de portecida de la portecida de la portecida de la portecida de la portecida de portecidad de la portecida de la portecida del portecidad de la portecida de la portecida de la portecida de la portecida de la portecidad de la portecidad de la portecidad del portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad del portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad del portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecidad de la portecida							farada tatas Branch Katab Maraka Katabaga Katab Maraka Katabaga Katab Maraka Katabaga Katab	An one source analysis of the analysis of the memory and memory an		
Minister of	Adust interate	198 p. 19 Mari (19 10) 14 (19)		national distance a Mina Malaria analysis distance distances and analysis distances and a statements	Na Rocard Include and Alex		An approximation of the second	An alora " approximation for a real field of the second second for a real field of the second second field of the second	Application carbons, no contains			
	-	The states	n de la Argani a la face de la Argani Male Nacionalita de la Argania de									
	address for a		remain ann a fhi dhi an 1 dha									

Questionnaire

Tuesday, Isruary 18, 2023 3 29-PM

Intent

Send on the questionnant to all Metalvolve members, then sing interactive sections with two or more storements, listedly, we would be able to tak to one member with a cache and one without. The top section "Questions we want," would be forwarded to Metalvolve members, "We can dig into some of these questions into the biointeraction is used as pode them in the "Suff we need to incom" sections.

- · Work to modify the questions to include examples or better descriptions or examples on technical areas
- Now does a user get data out of the system.

Questions we want to ask of the MetaArchive members: https://athuk.com/wks/joks-demon

- Preferred way to have the system inpest data? [these might need more explanation or examples]

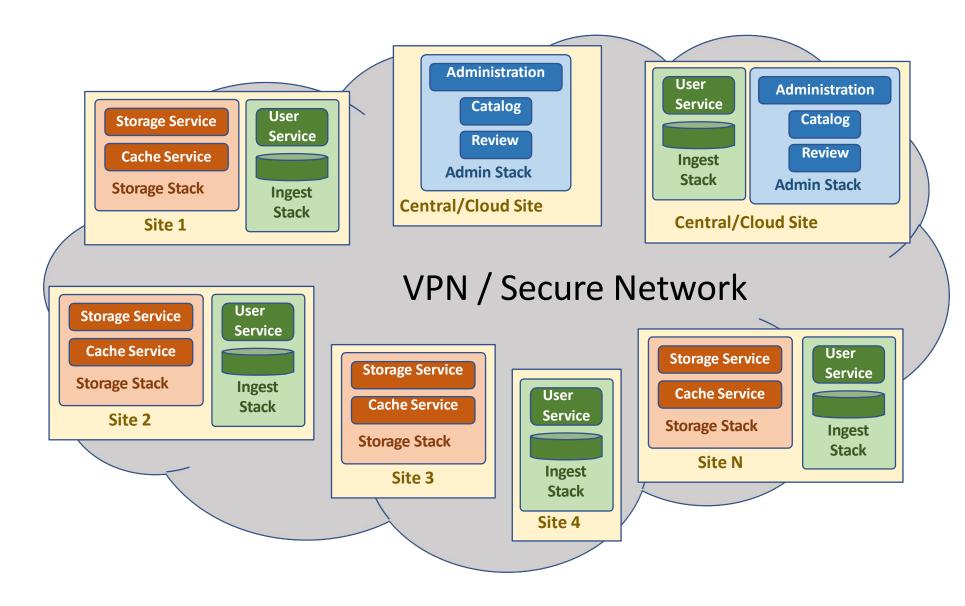
 TRV/VED
 - 0. 6852
 - o Web?
- What level of granularity is desired for access and sharing? (Deergies might help)
- What is holding you back from using the current capability more?
 If that didn't must, would use preserve more data into the system?
- · What kind of (descriptive) metadata do you enter to keep track of assets?
- o What could be done to make this easier?
- How large are the archive objects you want to submit to the storage environment?
 Min-stor? Max size? Average size?
 - How many such objects do you espect to submit?
 - As of April 2021, members and they expect (hope) to ingest 45-157 f8 within the next three years (2022-2025)
 We have less than 25 T8 of unique details the network now
- Inighest annual growth rate ever is about 4/18 year
- Now do you define where your data can be stored? (Which member caches?)

 Do you want to have more control on where your data is stored?
- Here the wave to chain more control on inferre poor place is stored?
 Here these is not currently controlled/specified by mentions. Target caches are accigned by central staff based on cache rapacity and geographic distributions to cool market are filled as to have doe capability?
- Should the login credentials from your organization be utilized for gaining access to the preservation system?
- Is there an asset iffetime?
- o . Hold Does this mean, are there are assets that do not require indefinite preservation or have a set retention period?
- Heave tell us a little bit about your current (Her local?) infrastructure.
 Her is their bit bit about your current (Her local?) infrastructure.
 Her is their bit bit bit about your current (Her local?) infrastructure.
 Statistic bit bit bit bit bit about your current (Her local?) infrastructure.
 Her is the information, but something to when when when we distribute the sorvey (or potentially this is a set of questions that should go in a different instructions grave).
- Amount and type of storage (might need explanation or examples)
- APS, SMB, S3, etc.
 Metwork connectivity to peer sites
- Over authentication
- LOCKSL Architemetica inte
- XW. Depending on what you'd like to know about currently workflows, another set of tools to ask about are ones for oreating Bagt large – Region, Exectly, DART (SUR) and tagkt py (CU).
- Mould you be interested in participating in a follow up interview with Respondent to further docum your workflows and needs?

• Responses:

- Difficult to Use
- Understaffed
- Automation
- Ease of Ingest
- Physical Shipment
- Wide variety of expertise, infrastructure, workflow

Architecture

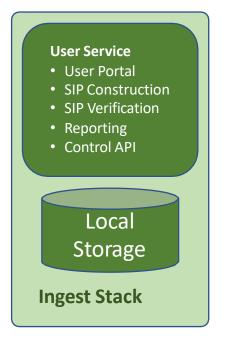


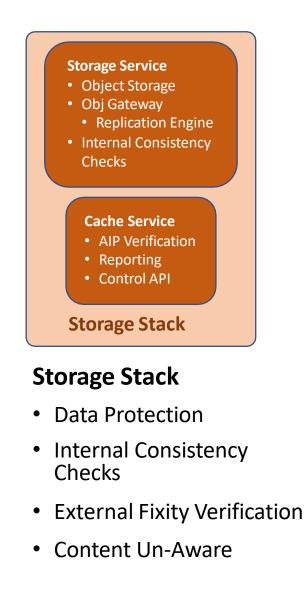
- Goals:
 - Easy
 - Automation
 - Increase Usage
- Distributed Architecture
 - Across Member Sites
- Modular Design
 - Well Defined Interfaces
- Scalable
- Secure
- Extensible

Components

Ingest Stack

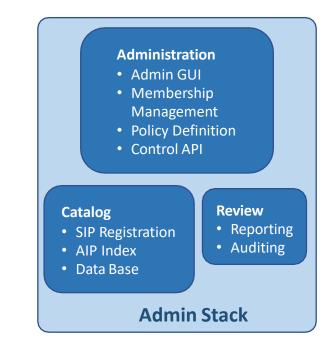
- Primary User Interface
- Local or Cloud
- Archive-Ready Packages and Native Assets





Admin Stack

- Limited Access
- Globally Unique Registration
- Additional Metadata
- Policies, Membership, Etc.



THANK YOU

Nathan Tallman Penn State University <u>ntt7@psu.edu</u> @ntallman@digipres.club Glenn Heinle KeeperTechnology glenn.heinle@keepertech.com

Library of Congress Designing Storage Architectures 2023