

Budget & Five Year Plan Work for Backup Collaboration Update for 9/22/2022 Backup Collaboration Steering Committee

TOTAL STARTUP FOR BACKUP COLLABORATION \$580,684 + \$35,885 = \$616,569

Total available terabytes = 85.2323 Total available gigabytes = 85232

TOTAL ANNUAL OPERATING COSTS FOR HOST SITES (not including staffing) = \$24,000 (for both backup and digitized materials archives)

LSTA RESERVE funds total: \$243,870

August 29, 2022 J. Klingbeil, C. Sommersmith, and V. Teal Lovely met to review the [draft Budget for Backup Collaboration](#). After reviewing the proposed budget, the group developed a list of questions to be discussed with the full group (which includes project administrators).

September 6, 2022 K. Hilbelink, A. Hoks, J. Klingbeil, K. Schwartz, C. Sommersmith, and V. Teal Lovely met to review the draft budget and discuss the questions from August 29.

1. Should the host systems charge for staffing to support the services and project management?
 - a. There is a conflict between keeping costs low so that systems can afford to participate in the backup solution and also the potential need to provide financial support to the host sites for project administration services.
 - i. *This is an issue to take up with the system directors of staff from the host sites.
 - b. If we oversized the current solution, could we get a smaller solution?
 - i. Yes. We could.
 - ii. The current system was sized as if all of the systems in the state would use it. At this time, 5 systems are using it and it is anticipated that 2 or 3 more would join. The solution is oversized for the 5 to 8 systems that will actually be on it and it would not be sustainable to replace it at the rate it is currently sized.
 - c. Is it possible to measure the raw storage before it is compressed? Is it possible to measure the data after it is compressed, backed up and stored?
 - i. If we use raw storage, it won't compare to other services, like Amazon, nor other variable costs.
 - ii. Raw storage is tangible and we know what that is in terms of bytes per system. But, this does not compare to pre and post-compression data and across different system types, etc.
 1. It's easy to get the amount of data that's protected, but it's not easy to get how much actual space the system is using.
 2. There is a report that comes with the Avamar client that is backend storage, but it is not working. This is the report that would be used to get the deduplicated storage.
 - a. In activity it shows what is being sent each night and how much. This is in the Avamar GUI and you click on

activity—you can see what is going back and forth. You can open it and see it. AH says these are the rounded numbers which are what he enters.

3. Having both raw data and compressed storage may help with cost comparisons.
 4. There may be no way to charge based on storage. There is data that is shared between systems, such as the Windows operating system at its core. Word documents may even share data. The shared data is assigned to the first system that adds it. You would have to know how much base data is deduplicated.
 - a. Would need to have a base cost, but even this is difficult.
 - b. Dividing the total costs just by number of systems is not fair either
 5. Could look at bytes processed. This would be difficult as well. Reports would need to be exported as a spreadsheet.
- iii. At some point the backup solution will also use the ECS (digitized materials archives).

Conclusions:

- We have a system that is already paid for.
- There are annual operating costs for host sites (not including staffing) (actual for LEAN, WI and estimated for SCLS) of \$24,000.
- We are using the costs of the current system to estimate replacement costs, but it's not necessarily the replacement cost.
- There may be no way to charge based on storage (use).
- Come up with an amount per system per year, rather than a fee based on use.
- Plan for 7 year replacement. Step systems toward the level that it actually costs by using DPI funds to pad the gaps. Move them toward sustainability over a period of time.
- Ask if DPI will match some amount.

Next steps:

- Talk to host site system directors about charging for staffing to support the services and project management.
- Ask DPI to cover a percentage of non staff costs until sustainability is reached.
- Kris, Andrew, Josh, Will, Kerri will define the services. They will ask the Cybersecurity Response Team for assistance.
- After service is defined, poll the systems to find out what they would be willing to pay for a quality system so that we could come up with what we have available as the total amount for a replacement system. (We also have LSTA funds in reserve.)
- The groups will continue to meet to complete the 5 year plan, which includes the budget.

Digitized Archives

There have also been meetings with the Recollection Wisconsin Digitized Materials Group to identify the system administration support for the archives. The group continues to meet to develop the pilot process.

Digitized Archives fees: It will be possible to create a flat rate per sized bucket, for example a 5 TB bucket. Digitized storage is less expensive because the IDPA is pulling the data that's being backed up, processing it and putting it into a backup that can be pulled easily. Most of this data will not be duplicated.