A. Organizational Infrastructure

A1. Governance & organizational viability

A1.1 - Repository has a mission statement that reflects a commitment to the long-term retention of, management of, and access to digital information. Yes

Evidence:

Chronopolis Mission Statement

Chronopolis is a digital preservation program for the preservation of long-lived digital data collections. It accomplishes this through the development and implementation of a preservation data grid and its supporting human, policy, and technological infrastructure. Chronopolis is intended as a model for valued digital collections with long-term impact from a variety of academic disciplines. The infrastructure is designed to be content-agnostic, to enable the ingest of collections of all types.

Chronopolis starts with the realization that partnership and collaboration among a specified group of like organizations is necessary to insure continuity and viability for the life-time of a collection. Thus it has been designed with multiple organizations providing the preservation infrastructure, tool and software development, and management. This is key not only to providing a robust preservation environment, but also creating a sustainable organizational model, not dependant on any one group.

To these ends, Chronopolis has been designed to:

• Provide long-term preservation of a wide variety of digital content;

• Operate a robust, grid-based storage environment for digital preservation;

• Maintain appropriate preservation metadata relevant to all aspects of the object lifecycle;

• Develop tools and services for digital preservation;

• Devote resources to investigating and planning for new technologies and services;

• Utilize and create community-based standards and systems.

A1.2 Repository has an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope. In development

Evidence:

Chronopolis has carried out tests within the Chronopolis network and also with external service providers to ensure that data currently housed in the Chronopolis system can be exported as needed. These tests have taken several specific forms:

1) Tests with the University of North Texas and the MetaArchive Cooperative. These tests have

focused on building appropriate DIP packages to transfer the data into non-SRB/iRODS based systems. This has been done using the BagIt protocol. These tests have demonstrated that data can be moved into LOCKSS-based storage systems, as well as a storage system which is neither SRB/iRODS or LOCKSS-based.

2) Tests with non-Chronopolis storage systems within the SDSC storage environment. Large portions of the data currently housed within the Chronopolis environment were previously housed elsewhere within the SDSC storage system, in different storage environments and under different software and management systems. This data has been moved in and out of Chronopolis several times. This has demonstrated the ease with which data could be exported out of the system into a generic storage environment (i.e. not one under control of SRB/iRODS or any of the Chronopolis tools) and used in many data center locations.

3) Data recovery tests between SDSC and NCAR storage systems. The Chronopolis team rebuilt a complete collection on a geographically remote node, to demonstrate that large amounts of data could be moved and verified upon receipt. During this test all of the Chronopolis monitoring tools were used to detect any changes or problems that might have occurred during the transfer. In addition, all objects were checksummed before and after the test for verification. All details of this transfer were cited and documented in detail. This document is available on the Chronopolis Sharepoint site.

These processes demonstrate the basic technological infrastructure which underlie the development of a formal contingency plan. The Chronopolis team will create a formal succession plan based on these technology underpinnings. The main issue that still needs to be addressed for this plan is finding non-Chronopolis service providers who are able to accept the quantity of data that is resident in the Chronopolis system. At the moment these discussions are in the formative stage.

It should also be noted that part of this plan has already been developed and is included in the Chronopolis Term Sheet, which can be found on the Chronopolis Sharepoint site. This document includes information about the rights and responsibilities of the Chronopolis storage providers as it pertains to the data within the system. This is targeted at the possibility of one of the storage providers leaving the Chronopolis partnership for any reason. The term sheet lays out the amount of time that is needed for this to happen (i.e. how much notification is needed) as well as what responsibilities the departing service provider has to the other partners.

A2. Organizational structure & staffing

A2.1 Repository has identified and established the duties that it needs to perform and has appointed staff with adequate skills and experience to fulfill these duties. Yes

Evidence:

The Chronopolis team has been designed and staffed based on the long history of service provision by the individual institutions in the network. Based on this experience, there are staff at each institution who are dedicated to the project, while sharing duties with the larger infrastructure of their organizations.

The staffing at each site can be broken down into two categories:

- General duties common to all sites: because all of the sites are maintaining the same software, middleware and equivalent hardware, there are designated system administration tasks that all must undertake. This includes management of the entire service stack, including storage and service hardware, networking equipment, SRB and storage maintenance, and related

software.

- Specific duties: each site also undertakes specific duties, based on local expertise. There are three categories of specialty:
    - ACE/Replication Monitor development: the UMIACS staff are responsible for writing and maintaining these software packages;
    - Portal development: staff at NCAR are responsible for developing the data portal;
    - Financial and organizational management, quality assurance, library services: staff at SDSC/UCSD are responsible for project management, overall repository management and coordination, marketing and publicity and assessment. The UCSD staff are also responsible for library services such as metadata development.

Several relevant documents can be found on the Chronopolis Sharepoint site:

- Organizational chart, with staff titles at each site;
- Organizational chart with specific staff members at each site;
- Resumes for staff at all sites.

A2.2 Repository has the appropriate number of staff to support all functions and services. Yes

Evidence:

Please see the answer to 2.1.

In addition, Chronopolis has been functioning as designed since spring of 2007. This design was informed and guided by the requirements of the NDIIPP grants which created the network. In addition, NDIIPP required regular reports on the progress of the network, demonstrating that it was meeting all goals. These reports are publicly available and can also be found on the Chronopolis Sharepoint site.

A2.3 Repository has an active professional development program in place that provides staff with skills and expertise development opportunities. Yes

Evidence:

Chronopolis functions inside academic environments, and the larger institutions and staff have professional development opportunities through their respective larger organizations. In the grant funding that Chronopolis has received thus far, money has been available for professional development in the form of conference and meeting attendance. Chronopolis staff have taken advantage of these funds and attended and presented on Chronopolis at all major digital preservation, repository and digital library conferences.

A3. Procedural accountability & policy framework

A3.1 Repository has defined its designated community(ies) and associated knowledge base(s) and has publicly accessible definitions and policies in place to dictate how its preservation service requirements will be met. Yes

Evidence:

Chronopolis does not define specific, genre-based designated communities. This is intentional and has been an informing factor in the infrastructure, software and service provisioning of the network since the beginning. The methods for data ingest and access have not been defined based on any specific type or format of data. Instead, Chronopolis is designed to accept data from any provider, assuming that they are willing to meet the requirements presented in the appropriate SLA and submission requirements.

A3.2 Repository has procedures and policies in place, and mechanisms for their review, update, and development as the repository grows and as technology and community practice evolve. Yes

Evidence:

The Chronopolis team is committed to technology reviews of its infrastructure every 24 months. This is done as a part of the larger organizations under which the network functions. This is also guided by typical hardware refresh and support contracts, which must be maintained.

A3.3 Repository maintains written policies that specify the nature of any legal permissions required to preserve digital content over time, and repository can demonstrate that these permissions have been acquired when needed. Yes

Evidence:

Chronopolis has several documents which address the legal permissions of data ingested into and maintained in the system. The primary documents are:

- Service Level Agreement, which must be signed by all data providers for every data collection. This is a publicly available document, which can be found on the Chronopolis Sharepoint site. This document is built upon SLAs that are used within the larger storage context in which Chronopolis is situated. This document includes language explaining Chronopolis' rights and responsibilities as it pertains to restricted content.

- Submission Agreement, which must be signed for all data transmissions. Designed to work in tandem with the SLA, the submission agreement must be signed by all data providers. It further highlights the rights and responsibilities that are appropriate to the Chronopolis team and those which are undertaken by the data providers.

It should be noted that Chronopolis only requests permission to make copies of digital content for the purposes of preservation. The system does not seek to make copies for access by anyone other than the specific data providers for the data content, nor to make copies that are to be re-used in any way by the Chronopolis system.

A3.4 Repository is committed to formal, periodic review and assessment to ensure responsiveness to technological developments and evolving requirements. Yes

Evidence:

Please see answer in section A3.2. The Chronopolis team is committed to ongoing review of all technology and hardware in use within its environment.

A3.5 Repository has policies and procedures to ensure that feedback from producers and users is sought and addressed over time. Yes

Evidence:

Chronopolis provides several formal mechanisms for data providers. These are available to all providers at any appropriate point in time:

- Chronopolis holds weekly meetings, in which staff of all service providers may participate.. Historically, there has been attendance by them when appropriate. Agendas for these meetings and notes of all discussions within them are sent every week to all data providers. These are archived on the Chronopolis Sharepoint site.

- Chronopolis maintains an active email list which contains all discussions regarding its services. This email list includes all data providers and they are free to comment and contribute to it. All meeting notices and notes are sent to the list, as well as requests for feedback on service changes and interruptions.

- Chronopolis team members have been creating a data portal, which allows the data providers to view the status of their data in the system. This has been built driven by feedback from the data providers. Included on the portal is a link to provide feedback or ask questions of the Chronopolis technical staff at any time. In the near future, the portal will be highly advertised to the data providers as the single point of contact with the Chronopolis staff. This will include all support and related questions.

A3.6 Repository has a documented history of the changes to its operations, procedures, software, and hardware that, where appropriate, is linked to relevant preservation strategies and describes potential effects on preserving digital content. Yes

Evidence:

All of the changes to its infrastructure, both planned and unplanned, are discussed within the weekly meetings and on the Chronopolis email list, noted in section A3.5. Notes and agendas from these meetings are kept in the Chronopolis Sharepoint site. The archives from the Chronopolis email list are also available on the Chronopolis Sharepoint site.

The Chronopolis preservation strategies are noted in section B3. All of the decisions and procedures

made for the system are done so in light of these strategies.

A3.7 Repository commits to transparency and accountability in all actions supporting the operation and management of the repository, especially those that affect the preservation of digital content over time. Yes

Evidence:

As already stated in sections A3.6 and A3.7, Chronopolis is committed to making all operations transparent.

A3.8 Repository commits to defining, collecting, tracking, and providing, on demand, its information integrity measurements. Yes

Evidence:

The integrity measurements of the data within Chronopolis form the core of the business enterprise. These are tracked and made available in several ways:

- Audit Control Environment (ACE): as noted on the ACE website (https://wiki.umiacs.umd.edu/adapt/index.php/Ace), ACE is a system that incorporates a new methodology to address the integrity of long term archives using rigorous cryptographic techniques. ACE continuously audits the contents of the various objects according to the policy set by the archive, and provides mechanisms for an independent third-party auditor to certify the integrity of any object.

- SRB Replication Monitor: The SRB Replication monitor is a simple webapp that watches registered directories and ensures that copies exist at designated mirrors. The monitor stores enough information to know if files have been removed from the master site and when the last time a file was seen. In addition any action that the webapp takes on files is logged. The monitor does NOT do any type of integrity checking, this is the responsibility of additional components. It should be noted that the SRB Replication Monitor will be replaced with a new tool as the transition to iRODS is completed.

  The team developing ACE and the Replication Monitor are part of the core technological staff members of Chronopolis, and the work within Chronopolis is based on interactions with these tools and its developers. All data from ACE and the Replication Monitor are freely available to the data providers whenever they desire.

- Data Portal: staff at NCAR have been developing a data portal, which synthesizes the information available separately from ACE and the Replication Monitor (and other systems within Chronopolis), and presents it in a clear, easy to understand view. The portal is intended to be used by the data providers and has been designed in consultation with them. The current URL for this portal is https://chron-mcat.ucar.edu:8443/chron/chronStatus.htm.

A3.9 Repository commits to a regular schedule of self-assessment and certification and, if certified, commits to notifying certifying bodies of operational changes that will change or nullify its certification status. Yes.

Evidence:

Chronopolis is committed to a regular and reasonable schedule of self-assessment and certification. Self-assessment (based on the TRAC recommendations) will be done every two years, and re-certification will be considered every five years.

A4. Financial sustainability

A4.1 Repository has short- and long-term business planning processes in place to sustain the repository over time. No

Evidence:

The Chronopolis team has been working on short term business planning as it relates to the organizations in which it is housed. This has taken the form of costing and charge models, identifying synergies with the larger organizations, and looking for new business collaboration opportunities. Because Chronopolis has been grant-funded via NDIIPP for the first generation of its life, this was the driving funding vehicle. It also determined which data providers and customers could be worked with. Chronopolis is now looking at opportunities outside of this initial group.

Chronopolis will be undertaking longer term planning over the next calendar year, and has committed to doing so in the latest cooperative agreement with the Library of Congress. This work will again be done in conjunction with the larger institutions in which Chronopolis resides.

A4.2 Repository has in place processes to review and adjust business plans at least annually. Yes

Evidence:

To the current date, as noted in A4.1, Chronopolis has been operating under a series of short term contracts. This has necessitated constant review of the Chronopolis services and the funding arrangements to provide them. At least yearly a new contract has been produced that reflects the appropriate changes.

As Chronopolis moves into a longer term funding model, processes will be put in place to maintain this review model.

A4.3 Repository's financial practices and procedures are transparent, compliant with relevant accounting standards and practices, and audited by third parties in accordance with territorial legal requirements. Yes

Evidence:

All financial practices and transactions are done through the business offices of SDSC, the UCSD Libraries, NCAR/UCAR and UMIACS. These are all based on official legal and financial documents from these organizations and are subject to all of the rules and regulations as required. These are available for audit as is deemed appropriate by all bodies which interact with these organizations. Chronopolis does not operate outside of these financial and business organizations.

A4.4 Repository has ongoing commitment to analyze and report on risk, benefit, investment, and expenditure (including assets, licenses, and liabilities). Yes

Evidence:

Because Chronopolis is managed by major research institutions, there are contractual and legal demands that must be met. There are regular reviews conducted by the business offices of all three Chronopolis partner sites with respect to all aspects of the project. This is a regular process that is managed by the Chronopolis project managers in conjunction with the financial business officers in the partner institutions.

A4.5 Repository commits to monitoring for and bridging gaps in funding. Yes

Evidence:

This is one of the core functions of the Chronopolis management team, including the PIs and project management team. Because Chronopolis is transitioning from a fully grant-based funding organization to a fee-for-service one, recognizing and bridging these gaps is, and will continue to be, one of the main efforts undertaken by the team.

A5. Contracts, licenses, & liabilities

A5.1 If repository manages, preserves, and/or provides access to digital materials on behalf of another organization, it has and maintains appropriate contracts or deposit agreements. Yes

Evidence:

Chronopolis has written Service Level Agreements and submission agreements which must be signed by all data providers.

A5.2 Repository contracts or deposit agreements must specify and transfer all necessary preservation rights, and those rights transferred must be documented. Yes

Evidence:

Chronopolis has written Service Level Agreements and submission agreements which must be signed by all data providers. As already noted, Chronopolis only requests permission to make copies of digital content for the purposes of preservation. The system does not seek to make copies for access by anyone other than the specific data providers for the data content, nor to make copies that are to be re-used in any way by the Chronopolis system.

A5.3 Repository has specified all appropriate aspects of acquisition, maintenance, access, and withdrawal in written agreements with contributors and other relevant parties. Yes

Evidence:

The Chronopolis Service Level Agreement and submission agreements contain all of the necessary documentation.

A5.4 Repository tracks and manages intellectual property rights and restrictions on use of repository content as required by deposit agreement, contract, or license. N/A

Evidence:

Chronopolis does not provide access to the data in its system to anyone other than the data providers. The rights and restrictions of data within the system are the responsibility of the data providers. Chronopolis does not add or modify these rights, other than to provide preservation level storage for the data.

A5.5 If repository ingests digital content with unclear ownership/rights, policies are in place to address liability and challenges to those rights. N/A

Evidence:

All questions of digital content ownership and rights are the responsibility of the data providers. Chronopolis does not assume responsibility for these objects. As noted in the Service Level Agreement and the Submission Requirements documents, all liability for objects remains with the data providers.

B. Digital Object Management

B.1 Ingest: Acquisition of Content

B1.1 Repository identifies properties it will preserve for digital objects. Yes

Evidence:

Chronopolis uses several methods to identify the properties it will preserve. These are laid out in multiple documents, including the Chronopolis Mission Statement, the Service Level Agreement and the Submission Requirements document provided to service providers.

Because Chronopolis is content and format agnostic, the driving concern is that content given by the data providers is preserved and made available back to them in the exact format. This is accomplished by the combination of tools used by Chronopolis, including SRB/iRODS and the ACE Monitor.

B1.2 Repository clearly specifies the information that needs to be associated with digital material at the time of its deposit (i.e., SIP). Yes

Evidence:

The Submission Requirements document and the Service Level Agreement together address the information that needs to be supplied at time of deposit. This is primarily of two types:

1) Information about the data provider, e.g. contacts, organizational information, etc.
2) Information about the data objects themselves, e.g. size of collection(s), number of files, file types, etc.

Chronopolis is moving to a strongly "manifest-based" environment wherein all data submissions must be accompanied by explicit manifests which contain metadata required for the preservation of the objects.

B1.3 Repository has mechanisms to authenticate the source of all materials. Yes

Evidence:

There are two relevant pieces of evidence for this question:

1) The Chronopolis ingest procedure is manual: there are humans in the process at all significant stages. In order to deposit data, a representative at the data provider contacts a Chronopolis system administrator who works to have everything setup properly and help with anything that is needed in the process. During ingest, the Chronopolis system administrator monitors progress, looking for errors using the tools noted below. At the completion of the process, the Chronopolis system administrator contacts the data provider representative and either confirms successful completion of the process, or notes errors that need to be corrected.

2) Chronopolis' ingest and replication procedure rigorously tracks all objects as they are ingested in all nodes of the system. At each significant event, technical metadata is captured and stored (often via system log files). In addition, each of the tools within the system's architecture is used to track the status and provenance of objects in the system:

- SRB/iRODS maintains all of the objects in the system. It provides logs on where in the storage

system the objects can be found, whether there are errors associated with them, etc.
- Replication Monitor: tracks each of the storage nodes and detects if there are missing objects in any of them. It provides logs of all findings and creates error alerts as desired.
- ACE Monitor maintains logs for each of the checksum processes it runs. This allows for system administrators to know if objects have changed inappropriately at any point in their storage life.

B1.4 Repository's ingest process verifies each submitted object (i.e., SIP) for completeness and correctness as specified in B1.2. Yes

Evidence:

The Chronopolis ingest process includes a thorough checking of each object submitted, including checksum verification and verification of all metadata items included in the manifest.

B1.5 Repository obtains sufficient physical control over the digital objects to preserve them. Yes

Evidence:

All objects ingested into the Chronopolis system come into a single access point, located at SDSC. Once objects reach this point, they are constantly managed, maintained and monitored throughout their presence within the system. As noted, verification of the validity of all objects occurs upon ingest. Also, Chronopolis maintains appropriate log files for the ingest and storage procedure to demonstrate physical control. This includes logs files from SRB/iRODS, the SRB Replication Monitor and ACE.

B1.6 Repository provides producer/contributor with appropriate responses at predefined points during the ingest processes. Yes

Evidence:

During the ingest process, Chronopolis staff work closely with the data providers. This includes discussion (usually via email) at all points to verify that data have arrived, the proper verifications have taken place, and that there are no issues. If there are issues, they are addressed and rectified as needed. As outlined in the Chronopolis Submission Requirements, the data provider receives a formal conformation that ingest has been successful.

B1.7 Repository can demonstrate when preservation responsibility is formally accepted for the contents of the submitted data objects (i.e., SIPs). Yes

Evidence:

As noted in section B1.6, the ingest process is iterative with the data provider and the log files generated during the process demonstrate the successful completion of the ingest. The formal ingest confirmation notification serves as the confirmation of acceptance. This notification includes a complete list of the objects ingested for that particular transaction.

B1.8 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (Ingest: content acquisition). Yes

Evidence:

Please see answers to the previous four questions.

B2. Ingest: Creation of the Archival Package

B2.1 Repository has an identifiable, written definition for each AIP or class of information preserved by the repository. Yes

Evidence:

Chronopolis is moving to a manifest-driven structure for referencing AIPs. There will be a single ingest object consisting of a package of arbitrary digital objects in the BagIt format. For this ingest object Chronopolis has a single AIP type, or class of information preserved. The only strict requirement for this AIP is a package manifest and at the data provider level an SLA. Each package ingested thus has a detailed manifest accompanying it that describes the objects contained within and their specific characteristics (identification, location within the collection, fixity). In addition to the manifest, for each package the data provider is encouraged to include content metadata in a separate structured file within the package which is included in the AIP. . The AIP also includes provider metadata harvested from the SLA.

This is stated within the Submissions Requirements document that data providers must sign.

Please see the Chronopolis System Description Document and the Chronopolis object model, which can be found on the project Sharepoint site.

B2.2 Repository has a definition of each AIP (or class) that is adequate to fit long-term preservation needs. Yes

Evidence:

In addition to the information noted in section B2.1, which describes the AIP that Chronopolis maintains, based on the original SIP, the Chronopolis system maintains a detailed log of all metadata that the system creates and uses to preserve objects in the system. This information can be found detailed in an excel spreadsheet on the project Sharepoint site.

The minimal data (as noted in B2.1) along with this detailed metadata creates an AIP within Chronopolis that is adequate for the preservation service that Chronopolis provides.

B2.3 Repository has a description of how AIPs are constructed from SIPs. Yes

Evidence:

As noted in section B2.1, Chronopolis is moving toward a manifest driven submission process. This creates a SIP which is minimally consistent for all objects. The process of ingesting content into Chronopolis is one of unpacking, i.e. taking data from the bags supplied by data providers and putting

it into the appropriate storage locations with the necessary pathnames and file structure. With this process of unpacking the bags, the AIP is a transformation / unpacking of the SIP. The content is then managed by the system with the information that Chronopolis adds to it. Chronopolis also adds information about the data provider that is provided in the SLA. The info.txt file in the bag provides data about the content of the object.

B2.4 Repository can demonstrate that all submitted objects (i.e., SIPs) are either accepted as whole or part of an eventual archival object (i.e., AIP), or otherwise disposed of in a recorded fashion. Yes

Evidence:

The ingest process for each SIP is separately monitored until completion. Chronopolis maintains detailed logs of all ingest and replication processes. The ingest process is not considered complete until all objects have been verified properly. This includes verifying they are located in the proper place within the submitted collection(s). To be ingested into the Chronopolis system, all submitted items must be assigned to a new or existing collection

B2.5 Repository has and uses a naming convention that generates visible, persistent, unique identifiers for all archived objects (i.e., AIPs). Yes

Evidence:

Chronopolis uses storage directory pathnames as unique identifiers. In this way uniqueness is insured during the time of ingest. If an attempt is made to write more than one object into the same directory path, a write error is produced and the process is not allowed without the manual intervention of a system administrator. The directory pathname includes /ChronopolisNode/dataProviderName/CollectionName/BagItPackageName/ as a precursor to the pathname of the object within the BagIt. In addition, as part of the Submission Requirements the data provider is mandated to provide unique identifiers (paths) to each file within a BagIt package. The data provider is additionally mandated to provide unique BagIt package names for each package in a collection, as well as unique collection names under their holdings. In this way the overall pathname within Chronopolis becomes unique. Chronopolis does not rename or modify the portion of the pathname or identifier originating from the SIPs as they become AIPs. All names (and also directory structures, etc.) from within the package remain the same as supplied by the data provider. It is the storage location within Chronopolis which is unique and is thus used as an identifier. Note that each Chronopolis node has a unique identifier which preempts the pathname ensuring that the same object residing on two Chronopolis nodes maintains a unique name within Chronopolis.

B2.6 If unique identifiers are associated with SIPs before ingest, the repository preserves the identifiers in a way that maintains a persistent association with the resultant archived object (e.g., AIP). Yes

Evidence:

A unique identifier is formally associated with the SIP during ingest but not before. The identifier is the location of the SIP and is thus ensured to be preserved with the AIP.

B2.7 Repository demonstrates that it has access to necessary tools and resources to establish authoritative semantic or technical context of the digital objects it contains (i.e., access to appropriate international Representation Information and format registries). N/A

Evidence:

At this point, Chronopolis does not have established relationships with external registries. This action is not part of the Chronopolis preservation plan (noted in section B3).

B2.8 Repository records/registers Representation Information (including formats) ingested. Yes

Evidence:

Within the manifest for each AIP the data provider maintains file type suffixes such as .doc, .jpg, etc. With this content metadata Chronopolis has some record of information formats. This is adequate to preserve objects in accordance with the Chronopolis preservation plan (noted in section B.3).

B2.9 Repository acquires preservation metadata (i.e., PDI) for its associated Content Information. Yes

Evidence:

Chronopolis maintains checksum information for each object in the system as well as file size and filename. These are maintained throughout the life of the object within the Chronopolis system.

B2.10 Repository has a documented process for testing understandability of the information content and bringing the information content up to the agreed level of understandability. N/A

Evidence:

This is not applicable per the Chronopolis Service Level Agreement and the preservation services that Chronopolis provides.

B2.11 Repository verifies each AIP for completeness and correctness at the point it is generated. Yes

Evidence:

As noted, Chronopolis maintains a manifest-driven system of ingest and preservation. The manifest represents the minimum requirement for an AIP. If it is incomplete the package will not ingest without errors. This allows for verification of all significant metadata characteristics as designed by the data provider. At all steps of the ingest and AIP creation process, verifications of the objects, comparing them to the appropriate manifests, occurs.

B2.12 Repository provides an independent mechanism for audit of the integrity of the repository collection/content. Yes

Evidence:

One of the key tools within the Chronopolis system is the Auditing Control Environment (ACE), a tool generated specifically for monitoring the integrity of the collections. Using ACE, Chronopolis maintains a regular schedule for auditing all objects in the collection. The ACE tool runs independently at each Chronopolis node. Resulting fixity data is compared between nodes and against the authoritative values. Authoritative values in turn have their own fixity values which are tokens stored outside the Chronopolis system and are regularly checked for completeness.

Full details of ACE can be found here:

https://wiki.umiacs.umd.edu/adapt/index.php/Ace

Note that the team which created ACE, the ADAPT group at UMIACS, is one of the core members of the Chronopolis team.

B2.13 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (AIP creation). Yes

Evidence:

Yes. All tools within the Chronopolis system, including SRB/iRODS, ACE and the Replication Monitor, provide detailed logs for all actions and processes within the system. These logs are under constant management of the Chronopolis system administrators and are captured and preserved as AIPs within Chronopolis.

B3. Preservation Planning

B3.1 Repository has documented preservation strategies. Yes

Evidence:

As noted in its mission statement, Chronopolis has a preservation strategy which targets the creation, maintenance and persistence of a preservation infrastructure, Specifically:

Chronopolis has been designed with multiple organizations providing the preservation infrastructure, tool and software development, and management. This is key not only to providing a robust preservation environment, but also creating a sustainable organizational model, not dependant on any one group.

To these ends, Chronopolis has been designed to:

• Provide long-term preservation of a wide variety of digital content;

• Operate a robust, grid-based storage environment for digital preservation;

• Maintain appropriate preservation metadata relevant to all aspects of the object lifecycle;

• Develop tools and services for digital preservation;

• Devote resources to investigating and planning for new technologies and services;

• Utilize and create community-based standards and systems.

It should be noted that the notion of format obsolescence is not at this point in time a concern of the Chronopolis system. Instead, this is taken to be the responsibility of the data providers who work with Chronopolis. It is explicitly stated in the SLA and Submission Requirements of the Chronopolis system that it will preserve the objects deposited in it in such a way that they can be transmitted back to the data providers in the exact form they were submitted.

B3.2 Repository has mechanisms in place for monitoring and notification when Representation Information (including formats) approaches obsolescence or is no longer viable. N/A

Evidence:

Chronopolis does not provide such mechanisms. At this point, all responsibility for format obsolescence is the responsibility of the individual data providers. The content information needed to provide monitoring and notification of format obsolescence is captured should Chronopolis decide to provide this service.

B3.3 Repository has mechanisms to change its preservation plans as a result of its monitoring activities. Yes

Evidence:

Because Chronopolis' mission is to provide a robust, appropriate preservation infrastructure, one of its key functions is to be flexible and able to change its preservation plans as a result of monitoring its system. The management and system administrators of the Chronopolis infrastructure have already taken steps in the past to implement improvements and refinements to its preservation plans based on the needs of the system.

B3.4 Repository can provide evidence of the effectiveness of its preservation planning. Yes

Evidence:

As already noted, Chronopolis' preservation strategy is based on planning and implementing an appropriate and robust hardware and software infrastructure to maintain large collections of data objects. The effectiveness of this strategy is judged on the ability of the Chronopolis system to ingest, maintain, and disseminate objects supplied by the data providers. Chronopolis has demonstrated all steps in this process:

1) Chronopolis has ingested 25+ terabytes of digital content from multiple sources. This content was successfully verified and registered within the system.
2) Chronopolis has undertaken several infrastructure migrations, including new versions of all software used in the system, installation and configuration of new software, installation and configuration of new hardware, and migration of digital objects among generations of hardware and software. At all stages the digital content was monitored and its preservation demonstrated using the Chronopolis tools.
3) Chronopolis has disseminated almost 10 terabytes of data back to multiple data providers. The data providers were able to receive this content and verify that it had not been changed and

was identical to the original content that was supplied to Chronopolis. This dissemination has been done using multiple tools and transfer mechanisms.

B4. Archival Storage & Preservation/Maintenance of AIPs

B4.1 Repository employs documented preservation strategies. Yes

Evidence:

Please reference the answers in section B3, particularly section B3.1 for this.

B4.2 Repository implements/responds to strategies for archival object (i.e., AIP) storage and migration. Yes

Evidence:

The two core tools within the Chronopolis system (SRB/iRODS and ACE) are both tools which can be modified to manage and preserve objects as circumstances change. The tools provide multiple ways to store, reference, ingest, and monitor objects. In addition, the Chronopolis system has been designed to be flexible enough to have a robust migration path as needed. This has already been demonstrated, as several major migrations have taken place. For example, the transition from SRB to iRODS represents a major process which Chronopolis has undertaken. Also, significant changes in the underlying hardware infrastructure at the Chronopolis sites have occurred with no impact on the validity and safety of the AIPs.

It should be noted again that file format migration is not part of the Chronopolis preservation strategy.

B4.3 Repository preserves the Content Information of archival objects (i.e., AIPs). Yes

Evidence:

The authoritative manifests for the objects within Chronopolis are the main tools for maintaining the content information. These manifests are preserved and managed as if they themselves were archival objects, i.e. replicated and tracked within the system.

As also noted in section B3, Chronopolis is able to preserve the content of the AIPs in the system and has created DIPs that can be given back to the appropriate data providers as needed.

B4.4 Repository actively monitors integrity of archival objects (i.e., AIPs). Yes

Evidence:

As noted in Section B2.12, Chronopolis runs ACE to provide integrity checking of all objects on a regular basis.

B4.5 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (Archival Storage). Yes

Evidence:

The log files of SRB/iRODS and ACE provide logs of all actions taken on objects within the Chronopolis system. These logs can be used to show what has happened to an object during its lifespan in the system. These are captured and stored as AIPs within Chronopolis.

B5. Information Management

B5.1 Repository articulates minimum metadata requirements to enable the designated community(ies) to discover and identify material of interest. Yes

Evidence:

As noted previously, Chronopolis defines its designated communities as the data providers who deposit content into the system. Because Chronopolis maintains the data in the exact directory structure designated by the data providers, and stores with it the metadata that was provided, the data providers can access and browse their data to discover and identify objects in their collections. Chronopolis provides several tools for data providers to access, browse, and review their holdings in the system. The most important of these are ACE and the data portal. Using these tools they can identify what data is in the system, when it was last checked, and its replication status. The data providers are responsible for knowing the characteristics of the content in the system.

B5.2 Repository captures or creates minimum descriptive metadata and ensures that it is associated with the archived object (i.e., AIP). N/A

Evidence:

The Chronopolis system does not depend on descriptive metadata for the purposes of identification. The objects in the system are identified by their specific pathnames (which includes their filenames). The minimum descriptive data as stated in the Submission Requirements document includes filename and fixity value for each file in a SIP. An optional package and/or collection level description can be included by the data provider. Since the unaltered SIP becomes the AIP, this information is captured. Furthermore the Chronopolis system, within the MCAT/iCAT services (part of SRB/iRODS), captures system level metadata on objects.

B5.3 Repository can demonstrate that referential integrity is created between all archived objects (i.e., AIPs) and associated descriptive information. N/A

Evidence:

This is not applicable to objects in the Chronopolis system.

B5.4 Repository can demonstrate that referential integrity is maintained between all archived objects (i.e., AIPs) and associated descriptive information. N/A

Evidence:

This is not applicable to objects in the Chronopolis system.

B6. Access Management

B6.1 Repository documents and communicates to its designated community(ies) what access and delivery options are available. Yes

Evidence:

Chronopolis' designated communities are the data providers who have deposited their content. Chronopolis is currently a dark archive, and as such, the data provider alone is authorized to access and retrieve its content. This information is communicated to all data providers at all stages of discussion, informally as well as in the SLA and submission agreements.

B6.2 Repository has implemented a policy for recording all access actions (includes requests, orders etc.) that meet the requirements of the repository and information producers/contributors. Yes

Evidence:

As previously noted, Chronopolis is currently a dark archive, and the concept of "access" is strictly defined. The system does not provide any kind of automatic or unmediated access avenues. To access their data, providers must contact Chronopolis system administrators who discuss their needs and provide the necessary tool to handle the transaction. Most commonly, this means the Chronopolis system administrator will send them the data.

To further refine this process, the Chronopolis team has created support emails and specific support website access (on the data portal) for data providers. When these tools are used, the interactions contained are archived on the Chronopolis Sharepoint site for future searching and verification.

B6.3 Repository ensures that agreements applicable to access conditions are adhered to. Yes

Evidence:

As noted, the nature of the Chronopolis system means that data providers do not have direct access to the data in the system. This is done at the system level and monitored using security and system administration procedures.

B6.4 Repository has documented and implemented access policies (authorization rules, authentication requirements) consistent with deposit agreements for stored objects. Yes

Evidence:

See Section B6.1, B6.2, & B6.3.

B6.5 Repository access management system fully implements access policy. Yes

Evidence:

See Section B6.1, B6.2, & B6.3.

B6.6 Repository logs all access management failures, and staff review inappropriate "access denial" incidents. Yes

Evidence:

As previously noted, Chronopolis relies on a number of monitoring tools which are running constantly. All attempts to access data, whether successful or not, are recorded and available for review. If there are incidents which are deemed to be inappropriate or questionable, they are reviewed and dealt with as needed.

B6.7 Repository can demonstrate that the process that generates the requested digital object(s) (i.e., DIP) is completed in relation to the request. Yes.

Evidence:

The dissemination process within the Chronopolis system is a manual one: all transactions occur with the explicit interactions of a Chronopolis system administrator with a representative from a data provider. This process is not automated. All DIPs include a checksum tied to the object as well as relevant metadata that was supplied with the object. This is verified with the data provider following dissemination.

B6.8 Repository can demonstrate that the process that generates the requested digital object(s) (i.e., DIP) is correct in relation to the request. Yes

Evidence:

Please see the answer to section B6.7.

B6.9 Repository demonstrates that all access requests result in a response of acceptance or rejection. Yes

Evidence:

Since all requests are manual, they are handled by the appropriate system staff. Acceptance or denial for specific requests are noted and recorded as needed.

B6.10 Repository enables the dissemination of authentic copies of the original or objects traceable to originals. Yes

Evidence:

Because of the robust monitoring and logging capabilities of the tools used in Chronopolis, a clear provenance statement can be demonstrated for all of the objects in the system. The provenance statement indicates source and collection for any given object. This insures that objects which are disseminated to the data providers exactly match those that were originally supplied.

C. Technologies, Technical Infrastructure, & Security

C1. System Infrastructure

C1.1 Repository functions on well-supported operating systems and other core infrastructural software. Yes

Evidence:

The Chronopolis infrastructure is based on well-established, mature hardware and software components. There are two relevant overarching principles driving this:

The three nodes in the Chronopolis system (SDSC, NCAR and UMIACS) are all resident in large data centers with a track record of decades of production support. This is done to help guarantee that the hardware infrastructure is placed in locations with reliable power, cooling, support, etc. All of the sites have staff monitoring the data center 24/7.

The hardware components at each site are actually different. This is done to insure that Chronopolis does not come to rely on any one vendor or technology platform, or become dependent on specific technology components which may change with time.

All of the storage within the Chronopolis system is on production level commercial vendors who provide industry standard service support for issues and problems. In addition, the sites maintain robust storage archival and software management tools. This includes appropriate levels of RAID for data security and redundancy for fault tolerance, as well as archival software such as SAM-QFS to manage disk and tape access as needed.

Layered on the hardware and software components described above, the SRB/iRODS middleware manages how objects are stored and shared among the sites. This layer provides both additional logging and tracking of objects within the system as well as object abstraction, so that objects can be referenced outside of any specific locale, if desired.

A complete list of the relevant infrastructure components at the individual nodes can be found on the Chronopolis Sharepoint site.

C1.2 Repository ensures that it has adequate hardware and software support for backup functionality sufficient for the repository's services and for the data held, e.g., metadata associated with access controls, repository main content. Yes

Evidence:

The technical staff at the individual Chronopolis sites manage issues related to storage of content at their local sites. They also work (individually and with system administrators from other Chronopolis sites) with the necessary associated metadata for objects. Over time, content is migrated as needed for technology refreshes, the need for more storage capacity, or equipment damage.

Individual nodes within Chronopolis do not provide their own local backups of data, because the geographic replication of data within Chronopolis obviates the need for such backups. It should be noted that with RAID techniques discussed earlier there is disk hardware fault tolerance, making data loss at an individual site much less likely. If a situation arises where data needs to be replaced

due to loss or corruption, a verified copy will be moved from another Chronopolis storage node. This process has been demonstrated on several occasions with tests and real world examples.

It should also be noted that each of the Chronopolis nodes does provide a backup of the local metadata databases, which store the technical metadata needed to track objects within the system. These metadata databases are unique to each node and thus are appropriate to be backed up locally.

The SRB/iRODS software layer, in conjunction with the Replication Monitor, enables a Chronopolis node to re-ingest any amount of content onto new storage equipment. Individual Chronopolis nodes have already demonstrated that a complete migration of objects in a node can be done, with fixity checks on the transferred data across nodes. The Chronopolis system has also demonstrated that it can move forward in hardware migrations as needed without damage or loss of data.

C1.3 Repository manages the number and location of copies of all digital objects. Yes

Evidence:

As noted, the use of SRB/iRODS and the Replication manager provide this core functionality for the Chronopolis system.

C1.4 Repository has mechanisms in place to ensure any/multiple copies of digital objects are synchronized. Yes

Evidence:

As noted, the use of SRB/iRODS and the Replication manager provide this core functionality for the Chronopolis system.

C1.5 Repository has effective mechanisms to detect bit corruption or loss. Yes

Evidence:

As noted in Section 2.12, one of the key tools within the Chronopolis system is the Auditing Control Environment (ACE), a tool generated specifically for monitoring the integrity of the collections. Using ACE, Chronopolis maintains a regular schedule for auditing all objects in the collection. The ACE tool runs independently at each Chronopolis node. Resulting fixity data is compared between nodes and against the authoritative values. Authoritative values in turn have their own fixity values which are tokens stored outside the Chronopolis system and are regularly checked for completeness.

C1.6 Repository reports to its administration all incidents of data corruption or loss, and steps taken to repair/replace corrupt or lost data. Yes

Evidence:

When the ACE tool detects any issues, it provides an alert to Chronopolis system administrators. After this point, any actions that need to take place are manually determined by administrators. If it is determined that there is data loss at a node, the needed data is recopied from either another Chronopolis storage node or from the original data provider, as appropriate. When a step such as this is taken, the issue is reported to the appropriate people, whether it is Chronopolis management (in the case of problems due to infrastructure concerns) and the original data provider (in the case of data loss or corruption that cannot be fixed within the system).

C1.7 Repository has defined processes for storage media and/or hardware change (e.g., refreshing, migration). Yes

Evidence:

The refresh and migration of the Chronopolis storage nodes are driven by two primary concerns: ensuring that the infrastructure is up to date and under appropriate maintenance, and ensuring that there is adequate storage space for the system's needs. Because most hardware vendors follow a 3-year maintenance window, this has been the default upgrade cycle for the storage nodes. In advance of this deadline, the management and system administrators begin discussion of what equipment needs to be upgraded or phased out and what options exist for upgrade. All discussions are held in the context of the system as a whole, and purchasing and upgrade decisions are made jointly.

C1.8 Repository has a documented change management process that identifies changes to critical processes that potentially affect the repository's ability to comply with its mandatory responsibilities. Yes

Evidence:

Chronopolis has implemented a change management service which maintains a record of all system changes, including planned and unplanned outages, upgrades and relevant system modifications. This change management service is maintained at SDSC, and is available via a website for viewing. All entries into the change management software are done centrally, with all incidents being first sent via email to the change management manager, who enters them into the system and keeps track as needed. This change management software can be demonstrated as needed.

C1.9 Repository has a process for testing the effect of critical changes to the system. Yes

Evidence:

All critical changes to the Chronopolis infrastructure are tested at the individual storage nodes before they are implemented. This includes updates to hardware, software, and the software management tools, including SRB/iRODS and ACE.

The Chronopolis team is in the process of developing a more robust testing environment which would provide a central way to track this testing. It is likely that this system will become part of the change management process in the near future.

C1.10 Repository has a process to react to the availability of new software security updates based on a risk-benefit assessment. Yes

Evidence:

Because the individual storage nodes in Chronopolis reside within larger data centers, the system leverages the security procedures present in these facilities. Each of the data centers has a robust security environment in place for the basic pieces of infrastructure. In addition, the Chronopolis system administrators maintain vigilance in monitoring the security of the data management components on top of the basic infrastructure. Chronopolis has produced a security document which guides the decisions made by the organizations within the system. This document can be found on the Chronopolis Sharepoint site.

C2. Appropriate Technologies

C2.1 Repository has hardware technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when hardware technology changes are needed. Yes

Evidence:

Chronopolis team members spend significant time each week discussing and reviewing the hardware technologies used in the system. A significant section of the weekly Chronopolis staff meeting is devoted to these concerns as needed. In addition, the Chronopolis team maintains contact with the data providers to ensure that the system is meeting their expectations. Meeting minutes can be found on the Chronopolis Sharepoint site.

C2.2 Repository has software technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when software technology changes are needed. Yes

Evidence:

Chronopolis team members spend significant time each week discussing and reviewing the software technologies used in the system. A significant section of the weekly Chronopolis staff meeting is devoted to these concerns as needed. In addition, the Chronopolis team maintains contact with the data providers to ensure that the software system components are meeting their expectations. Meeting minutes can be found on the Chronopolis Sharepoint site.

C3. Security

C3.1 Repository maintains a systematic analysis of such factors as data, systems, personnel, physical plant, and security needs. Yes

Evidence:

Chronopolis has produced an extensive document, the "Risk Analysis & Risk Assessment Security Standard," which outlines in detail all of the relevant issues. This document is available on the Chronopolis Sharepoint site. It is intended to be a living document, which is modified as needed.

C3.2 Repository has implemented controls to adequately address each of the defined security needs. Yes

Evidence:

Please see section C3.1.

C3.3 Repository staff have delineated roles, responsibilities, and authorizations related to implementing changes within the system. Yes

Evidence:

Because of Chronopolis' nature as a dark archive, there are relatively few people who have authority to make changes. The roles of these people are clearly delineated within the system (as noted in the Chronopolis organizational chart on the Sharepoint site). In addition, the Chronopolis staff members only have responsibility for the infrastructure at their own local site. This provides an additional security safeguard.

C3.4 Repository has suitable written disaster preparedness and recovery plan(s), including at least one off-site backup of all preserved information together with an off-site copy of the recovery plan(s). Yes

Evidence:

Chronopolis has a robust and complete disaster recovery plan. It can be found on the Sharepoint site.