Delivery
Component and Connector View

Document number…… SKA-TEL-SDP-0000013
Document Type…….. REP
Revision……..…… 05
Authors…….…….…… R. Simmonds, S. Goliath, P. Wortmann, K. Kirkham, S. Sanchez
Release Date…… 2018-04-23
Document Classification…… Unrestricted
Status……Released

Lead Author	**Designation**	**Affiliation**
R. Simmonds	DELIV lead	UCT
Signature & Date:	RWJSimmonds RWJSimmonds (Apr 24, 2018)	

Released by	**Designation**	**Affiliation**
Paul Alexander	SDP Project Lead	University of Cambridge
Signature & Date:	Paul Alexander Paul Alexander (Apr 24, 2018)	

Table of Contents

List of Abbreviations

AAAI	Authorization, Access, Authentication and Identification
DOS	Denial of Service
HPSO	High Priority Science Objective
IAM	Identity and Access Management
I/O	input/output
IVOA	International Virtual Observatory Alliance
PCB	Process Control Block
SDP	Science Data Processor
SIA	IVOA Simple Image Access protocol
SKA	Square Kilometre Array
SRC	SKA Regional Centre
SSA	IVOA Simple Spectral Access
TAP	IVOA Table Access Protocol
URI	Uniform Resource Identifier
WAN	Wide Area Network

1. Primary Representation

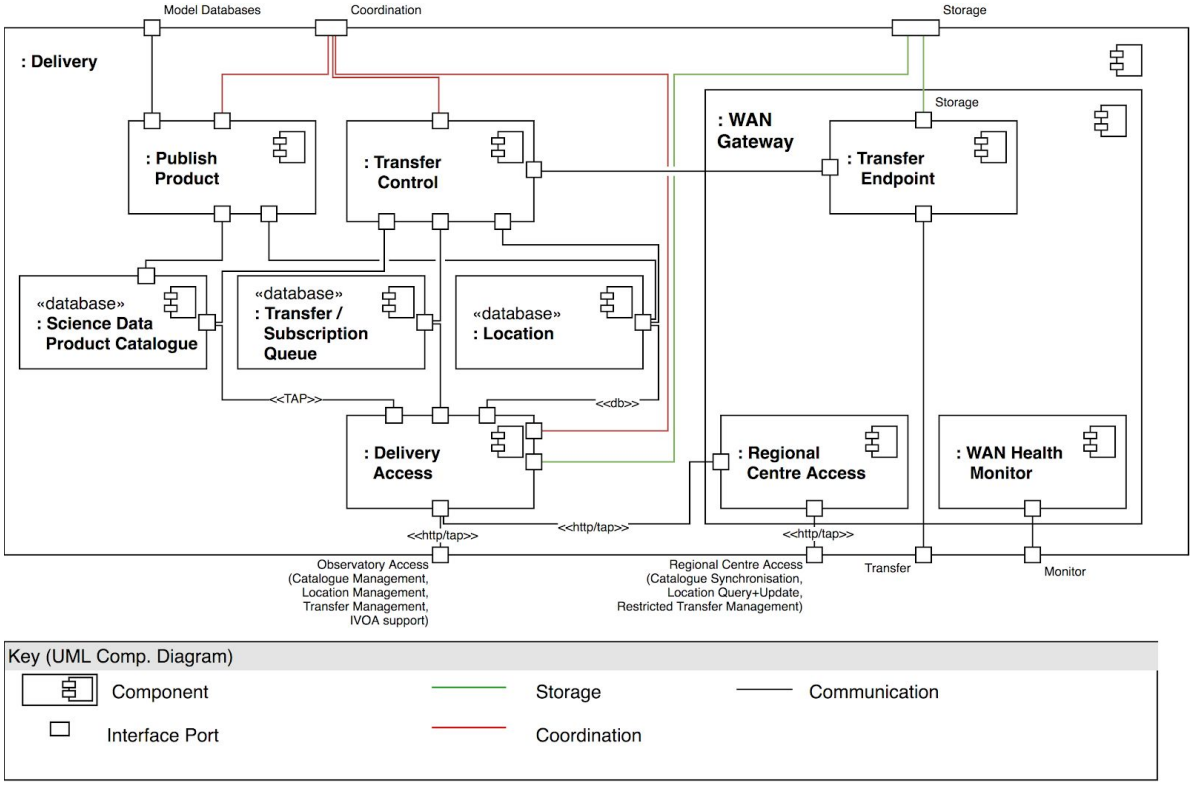

Figure 1: Delivery Primary Representation. This shows the internals of the Delivery component responsible for publishing the existence of data products and transferring them to SKA Regional Centres.

The diagram above shows the component and connector view of the Delivery component as it will be deployed at the SDP sites. The purpose of the Delivery system is to make data products available outside of the SDP sites. To do this, it publishes details of new data products and provides components to manage and perform product transfers. Publishing Products is achieved by adding entries to the Science Data Product Catalogue and creating an entry in the Location component that tracks instances of products that, as transfers occur, will be replicated at SKA Regional Centres (SRCs).

The Delivery System is expected to transfer approximately 1 PB of data per day, so hardware and network interfaces to the Wide Area Network (WAN) used for these transfers must be capable of moving data at close to 100Gb/s. According to the System Sizing Document [RD03], High Priority Science Objective (HPSO) data products will be generated at a rate of 25Gbits/s in LOW and 9Gbits/s in MID. The System Sizing Document also contains estimates that the size of HPSO imaging data products will range from few TBs to approximately one hundred TBs. This means that ten thousand entries would be added to the Science Data Product Catalogue each year relating to imaging data products. The size of non-imaging products, and consequently the number of entries in the Science Data

Product Catalogue for these, will depend heavily on how they will be packaged, which is still not fully determined. This could be as large as two million entries per year, though this number could be decreased considerably if products are packaged corresponding to longer collection times.

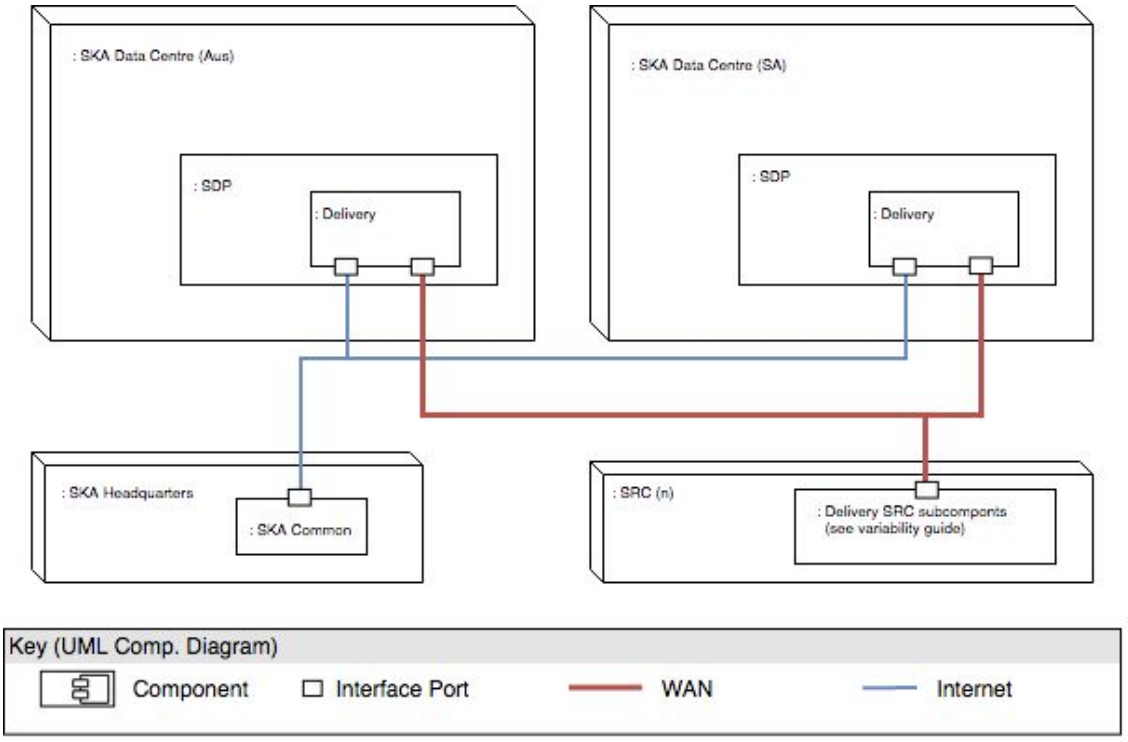

Figure 2: Delivery deployment view showing the primary components deployed at the two SDP sites, and deployed at the SRC sites to receive the data products and the replicated Science Data Product Catalogue. It also shows how the SKA Headquarters can access the Delivery system to manage data distribution policies and access IVOA services.

Referring back to Figure 1, the WAN Gateway component is exposed to the WAN network that links the SDP sites to other sites that will receive SKA data, notably the SRCs. It will accept inputs from these external sites through the Regional Centre Access component. The Delivery components that are not in the WAN gateway will only accept requests from the WAN that pass through the Regional Centre Access component.

The WAN Gateway transfers data between sites it connects to using the Transfer Endpoint component. The WAN Gateway also hosts a WAN Health Monitor component that is used to perform active monitoring of the WAN. As they perform different activities and have different security requirements, it is expected that the Transfer Endpoint, Regional Centre Access and WAN Health Monitoring will all be hosted on separate servers.

Data relating to the transfers is stored in the Transfer / Subscription Queue component and is accessed by the Transfer Control component that initiates the SDP Data Product transfers. The Science Data Product Catalogue holds information describing each data

product created by the SDP. The Science Data Product Catalogue is created at the SDP sites and replicated between the two SDP sites and out to sites in other regions. It is likely that each SRC will want a copy of the Science Data Product Catalogue to enable fast searches to be performed, though since it is likely to hold proprietary information, it should only be replicated to sites that can conform to the SKA data security policies.

The Publish Product component is used at SDP sites to add new entries to the Science Data Product Catalogue as new SDP Data Products are created at the SDP sites.

The Location Service that is used to track the location of SDP Data Products that have been transferred to SRCs. This only holds keys that reference data products, with a list of sites that hold the products.

The Delivery Access component allows Observatory staff to manage the data transfer policies by adding and removing subscriptions. It also provides access to IVOA services to Observatory staff. It also accepts authenticated requests from the Regional Centre Access component to a restricted set of services such as Product Request.

2. Element Catalogue

2.1 Elements and Their Properties

2.1.1 Publish Products

Used to update the Science Data Product Catalogue when new products are created. It subscribes for updates from the Execution Control operational system level component [AD02] for information about new products and adds this information to the Science Data Product Catalogue as it is made available. This information is received from a message queue which has the information added as the execution of that PCB completes. At this point it accesses the Model Databases to gather additional information needed to create new catalogue entries. Details of what is collected to create an Science Data Product Catalogue entry can be found in the SDP Data Model document [AD07]. In addition each entry is given a unique key which is computed from the scheduling block / processing block / data product id. Also, an entry is added to the Location database to indicate that the data product is available from this SDP site and to prepare the Location database for updates when the data product is replicated to other sites.

Reliability: Moderate, capable of reproducing catalogue entries when interrupted by failure in this service or dependent services. Capable of retrying operations when dependent services fail by reading from persistent message queues. Availability will not affect transfer elements.
Performance: Create Science Data Product Catalogue entries within minutes of the arrival of a new Process Control Block indicating that a workflow has completed.

Security: Authorised access is by SDP administrators or operators. Deliberate or inadvertent Denial of Service (DOS) attack will not interfere with other Delivery or SDP elements. This component is isolated from the external network.
Resources: Modest compute and I/O requirements, redundant hardware and network available for fail-over, redundant container deployment.

2.1.2 Science Data Product Catalogue

This catalogue identifies each of the SDP Data Products that is created at an SDP site. It includes all of the metadata that is needed to find a particular SDP Data Product. It does not contain all data items that may be needed for re-processing of SDP Data Products as individual objects. However, since snapshots of the Science Data Model are stored as SDP Data Products, additional information for data re-processing can be extracted from these snapshots at the SRCs. The metadata items used to create entries Science Data Product Catalogue is described in the SDP Data Model document [AD07]. Each entry also has a unique key used to cross reference entries in the Science Data Product Catalogue with entries in the Location database.

Reliability: Moderate availability. It should be able to recover from data loss and corruption using a backup.
Performance: Low I/O. Write new catalogue entries within few minutes. Answer catalogue queries within seconds.
Security: Only directly accessible from the SDP site by authorised users. Replication passes through the Regional Centre Access component that limits access to authorised sites.
Resources: Medium-large capacity storage with moderate transaction capability. Modest compute requirements. Redundant hardware and network available for fail-over.

2.1.3 Location

This is used to store the locations of SDP Data Products. This provides a list of sites a SDP Data Product is available from, referenced by the unique key for that SDP Data Product in the Science Data Product Catalogue. The construction of the URI to access a product at a site is performed by the Transfer Control at a site, so this does not need to be stored. Each key represents a single data product that is referenced in the SPDC.
Reliability: Moderate availability. It should be able of recovering from data loss and restoring corrupted data. Mechanisms to check the consistency of the information would be beneficial in case of update failures, though missing information will not cause the Delivery system cease working, it could just result in some additional requests from SRCs for data products that have already been delivered to other SRCs.
Performance: Low I/O. Write new database entries / Answer database queries within seconds.
Security: Only accessible through other Delivery components.
Resources: Medium-large capacity storage with low transaction overhead. Modest compute requirements. Redundant hardware and network available for fail-over.

2.1.4 Transfer / Subscription Queue

The Transfer / Subscription Queue is the data store used by Transfer Control to hold its state. The state includes the priority queue of the IDs for SDP Data Products for which transfers have been requested and a queue of handles to SDP Data Products that are ready to be transferred. The Transfer / Subscription Queue also holds the subscriptions that are used to create new transfer requests when new SDP Data Products are registered at a site. It also holds state used to place limits on the Transfer Control, such as number of file transfers active at any time, as well as statistics used to track the amount of data that is transferred to each site.

Reliability: High availability. Low availability will affect transfer control and and may delay data transfers.
Performance: Moderate transaction I/O. Enough to provide a responsive communication with Transfer Control.
Security: Only accessible through other Delivery Components.
Resources: Small capacity storage. Modest compute and transaction I/O requirements. Redundant hardware and network available for fail-over.

2.1.5 Transfer Control

This manages the transfers to and from a site. For transfers from a site it is responsible for managing the queued data product transfer requests. It does this by picking an entry from the request queue held in the Transfer / Subscription Queue component and requesting a handle to all the files associated with this product from the Buffer through the coordination interface. When a response is received to say that a file is ready to transfer, it queues ready to be transferred. Files will be passed to the Transfer Endpoint based on configured limits on the number of file transfers to keep active to any destination at any time. Transfer Control informs the Buffer though the coordination interface when it has finished transferring a particular file, so buffers can be freed.

The Transfer Control component also implements the subscription service. To facilitate data transfer, this compares subscriptions (rules for what data should be transferred to other sites) with available data products and past transfers, and creates new transfer requests for the data products which are now available. For example a subscription may state that all data associated with a particular project should be delivered to a particular SRC. Therefore as files relating to that project are added to the Science Data Product Catalogue, new file transfer requests are created and added to the request queue.

Reliability: Moderate, with the capability to recover from interrupted transfers. Availability will not affect Transfer Endpoint component.
Performance: Enough to generate new Data Product transfer requests from subscriptions and identify the data products with highest priority to be transferred when there is available capacity on the outgoing WAN link.
Security: Authorised access is by SDP administrators or operators to the control interface.

Resources: Modest compute and I/O requirements, redundant hardware and network available for fail-over, redundant container deployment.

2.1.6 Transfer Endpoint

Called when a product is ready to be delivered to setup a transfer with a specific remote endpoint at an SRC. It is called with the handles / URIs for the local files and the URIs that they will be delivered with to the external endpoint, which will be either an SRC or an observatory site such as the SKA Headquarters.

Reliability: Transfers should be possible while the Buffer holding data to be transferred is operational.
Performance: Requires Storage read and network write performance sufficient to fully utilise available network connections. Performance will downgrade when hardware fails. Nominal performance should be sufficient to keep up with products produced during normal telescope operations.
Security: Deliberate or inadvertent DOS attack will not interfere with other Delivery or SDP elements. Breach of this platform must not lead to access to the rest of the SDP or to access to other Transfer Endpoints, at SRCs for example.
Resources: Storage implemented on horizontally scalable fast media with tuned network endpoints.

2.1.7 Delivery Access

This provides a web-based user interface that is used for managing data transfers and for SKA staff to access IVOA services. The management of data transfers includes being able to request particular products to be transferred, or to insert subscriptions that provide rules for which products should be transferred to which site, based on metadata published about about the products in the Science Data Product Catalogue.

The IVOA Services accessible through this component are the Simple Spectral Access (SSA) [AD06], the Simple Image Access (SIA) [AD04], the Table Access Protocol (TAP) [AD03], and the DataLink [AD05] protocol. The SSA and SIA services will access the Science Data Product Catalogue via the TAP service, which will directly query the Science Data Product Catalogue. The DataLink protocol offers a binding mechanism to represent the logical links between metadata in the Science Data Product Catalogue, and data available to researchers at SRCs.

Reliability: Moderate. Delivery can keep working without this, but there will no longer be any ability to add new subscriptions or for the Observatory or SRC operators to monitor scheduled transfers. Failure will also prevent the Science Data Catalogue from being accessible to the Observatory or for new entries to get replicated to SRCs.
Performance: Capable of providing interactive response to queries. Catalogue queries should be responded to within a small number of seconds.
Security: Deliberate or inadvertent DOS attack could will limit ability to access service.
Resources: Modest compute requirements.

2.1.8 Regional Centre Access

This component provides the interfaces through which SRCs communicate with the SDP. It provides access to a subset of the services provided by the Delivery Access component. It provides catalogue synchronisation output which sends updates to the Science Data Product Catalogue to the SRCs and provides synchronisation input in the case this component is deployed at an SRC. Finally it provides the interface for transfer management for the SRCs.

Reliability: Low. Failures will prevent some interactions with SRCs, but will not stop other Delivery operations.
Performance: Capable of handling interactive use from SRC operators and replication traffic from the Science Data Product Catalogue.
Security: Deliberate or inadvertent DOS will limit interaction from SRCs and Science Data Product replication.
Resources: Modest compute and I/O requirements.

2.1.9 WAN Health Monitoring

Component that performs active monitoring of the WAN and provides an interface to historical monitoring information. This should at least collect available bandwidth and latency information. PerfSonar is an example of system that could be used for this. Having this within the WAN Gateway will ensure that testing is performanced on the “last mile” link to the SDP.

Reliability: Moderate. Failure will not prevent any other Delivery component from operating. However, keeping a full record of network health is
Performance: Needs to be able to run network at WAN speed for bandwidth testing.
Security: Moderate risk. Ideally historical data will be kept off-site so can be installed if compromised. However, a DOS attack could use all the available bandwidth on the WAN
Resources: Compute server and network interfaces capable of saturating the WAN network.

2.2 Element Interfaces

2.2.1 Delivery External Interfaces

Coordination
This interface is used by three Delivery components to exchange coordination information with the Execution Control and Buffer system level elements.

Coordination: Publish Product
Used to tell Publish Product that new metadata is available that might lead to an update of the Science Product Catalogue. Used to return health information to the Execution Control system element via the Delivery Control element.

Coordination: Transfer Control
Used to request files to be transferred from the buffer.
Used to return health information to the Execution Control system element via the Data Control element.

Coordination: Delivery Access
Used to request Data Products for use by IVOA services.

Storage
This interface is used by two Delivery components to access data products stored in the SDP buffer.

Storage: Transfer Endpoint
Used by the Transfer Endpoint to read data that it is transferring to a remote endpoint.

Storage: Delivery Access
Used by the IVOA services that need to access data products.

2.2.2 SDP External Interfaces relating to Delivery

Regional Centre Access
Interface for access to the Delivery component from SRCs.

Regional Centre Access : Catalogue Replication
Used to replicate the contents of the Science Data Product Catalogue to remote sites. It is likely that all SRCs would want a replica of this to enable fast searching, though it is possible that a small number of replicas managed by the Observatory will suffice.

Regional Centre Access : Location Query and Update
Used to make data product location information available, so that the sites holding replicas of particular products can be found, and for updating the locations of data products.

Regional Centre Access : Restricted Transfer Management
This provides ways for SKA Regional Centres to request data products and obtain monitoring information on transfers that have already been requested. Access to this interface and what can be requested over it will be set by a SDP operations policy, when it arises.

Transfer
The bulk data transfer interface used for sending and receiving data products. GridFTP is a suitable protocol for use on this interface with existing research networks.

Monitor

Used by the WAN Health Monitor to test the network by contacting monitoring services located at remote sites.

2.3 Element Behavior

2.3.1 Delivery Interactions

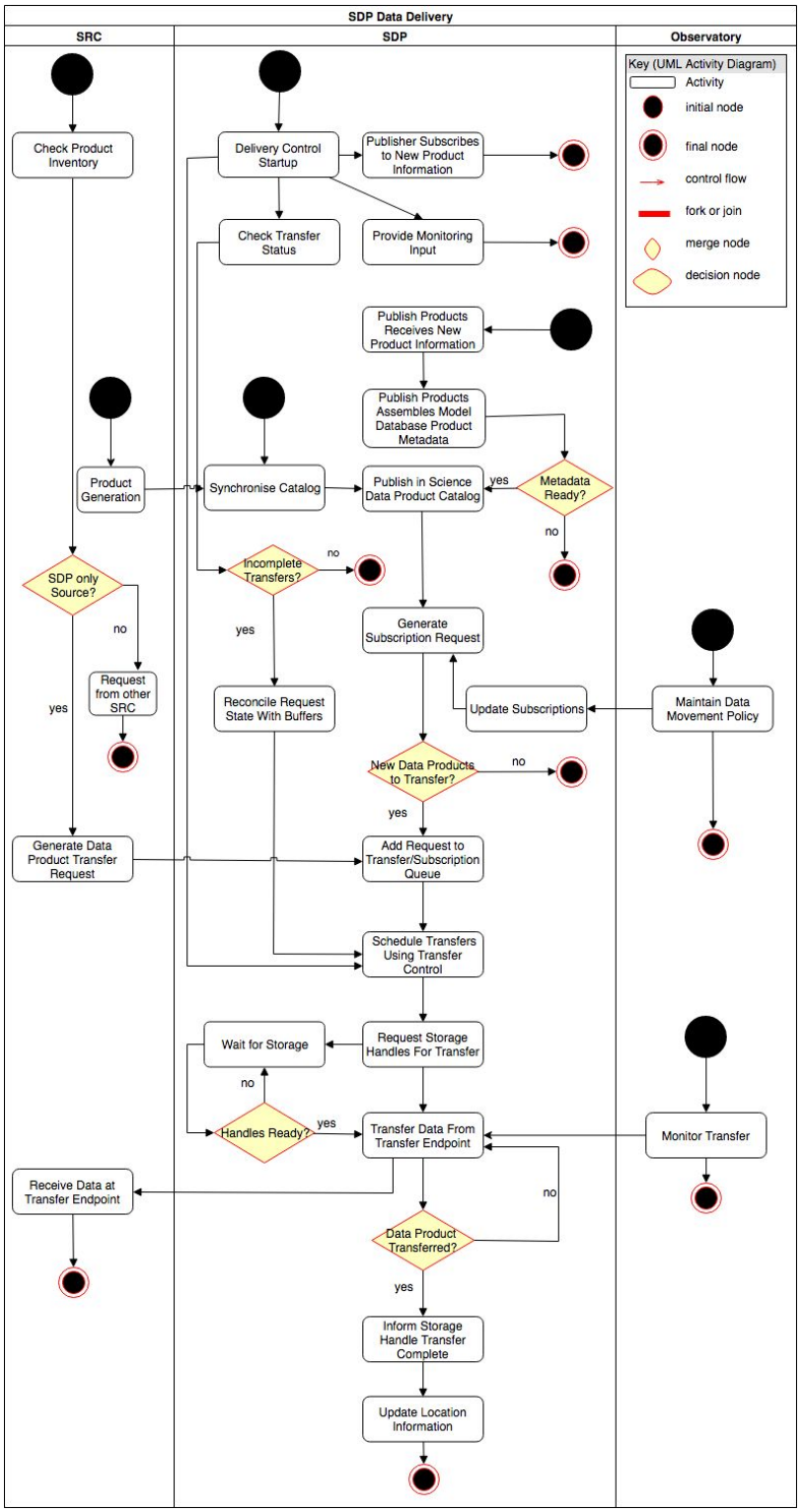

Figure 3: Activity diagram showing interactions between Delivery System and other components for its data transfer activities.

The left lane shows what an SRC should do when it wants a data product that it does not have. It should first check using the Location database if the data product is available from and SRC and request it from there if policy allows. Operations staff at the SRCs can also monitor the progress of transfers that have been requested for their site. The right lane shows that the Observatory staff can maintain a data transfer policy by adding and removing subscriptions and can monitor transfers. This centre lane shows the actions of the Delivery component on startup, and when it receives a Process Control Block to indicate that new data products may have been created.

2.3.2 Science Data Product Catalogue Entry Creation

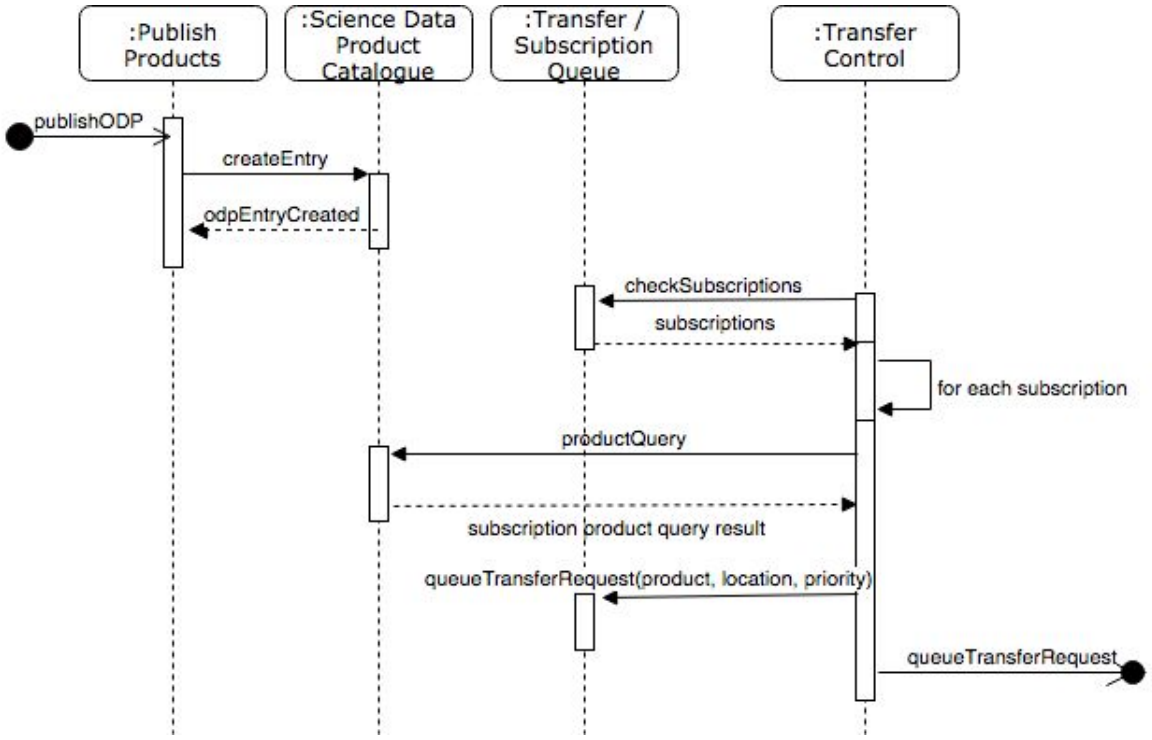

Figure 4: Sequence Diagram showing Science Data Product Catalogue Entry Creation

Catalogue entry creation occurs during normal SDP operations. The Delivery element utilizes the subscriber role in publish-subscribe behaviour for discovering information about new metadata for SDP Data Products. New catalogue entries may result in a subscription policy being fulfilled via a subsequently-generated request, and are replicated with eventual consistency to SRCs.

2.3.3 Data Subscriptions

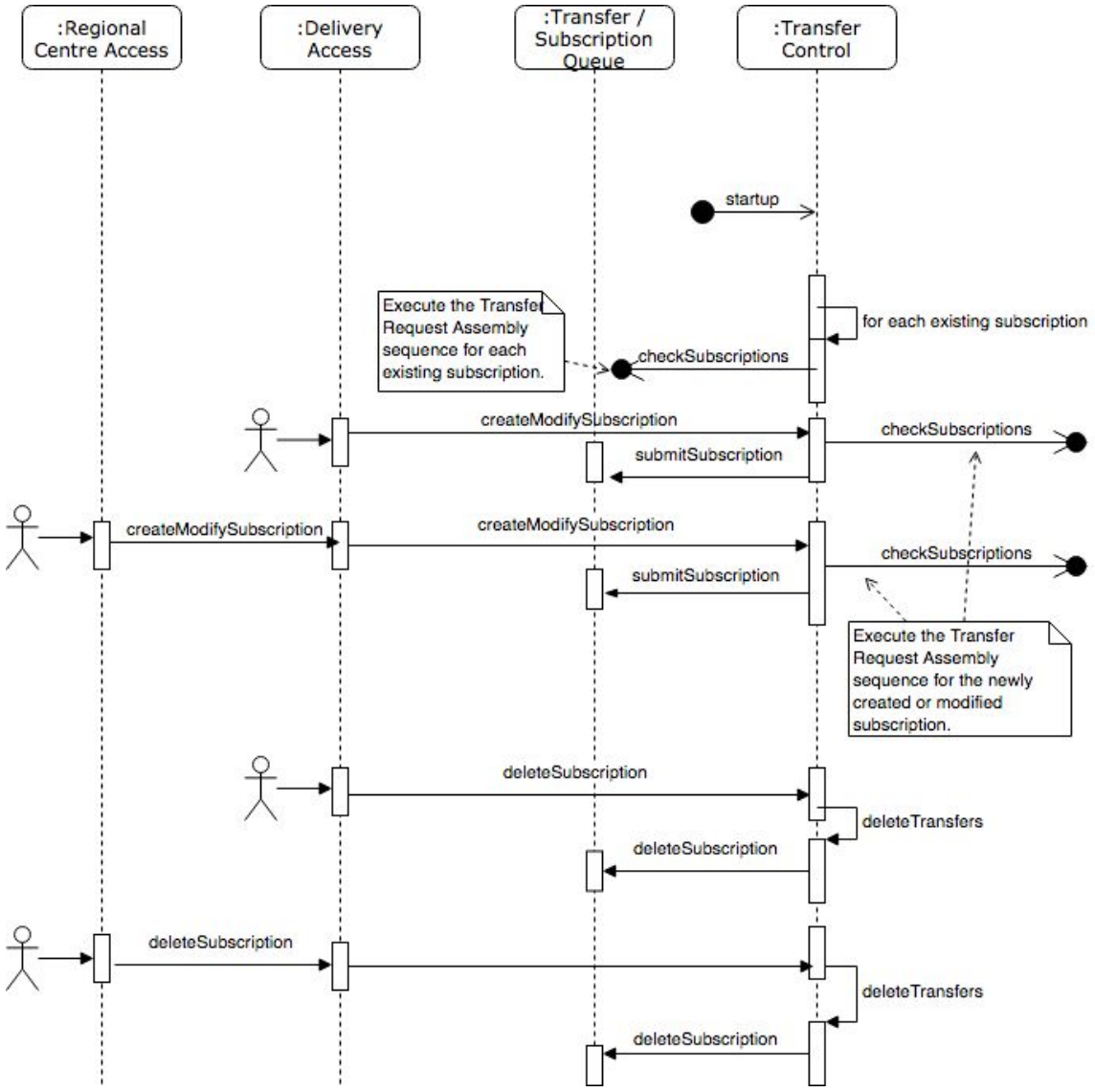

Figure 5: Data subscription lifecycle

The Subscription lifecycle, as influenced from the Observatory, the SDPs, or any of the Regional Centres, is shown in Figure 5. Each checkSubscriptions will trigger the actions in Figure 4.

Transfer Control will reconcile the state of existing transfer requests and SDP Data Product locations on startup, because shutdown may have been irregular. Any discrepancy will be resolved with a new transfer request.

2.3.4 Data Transfer

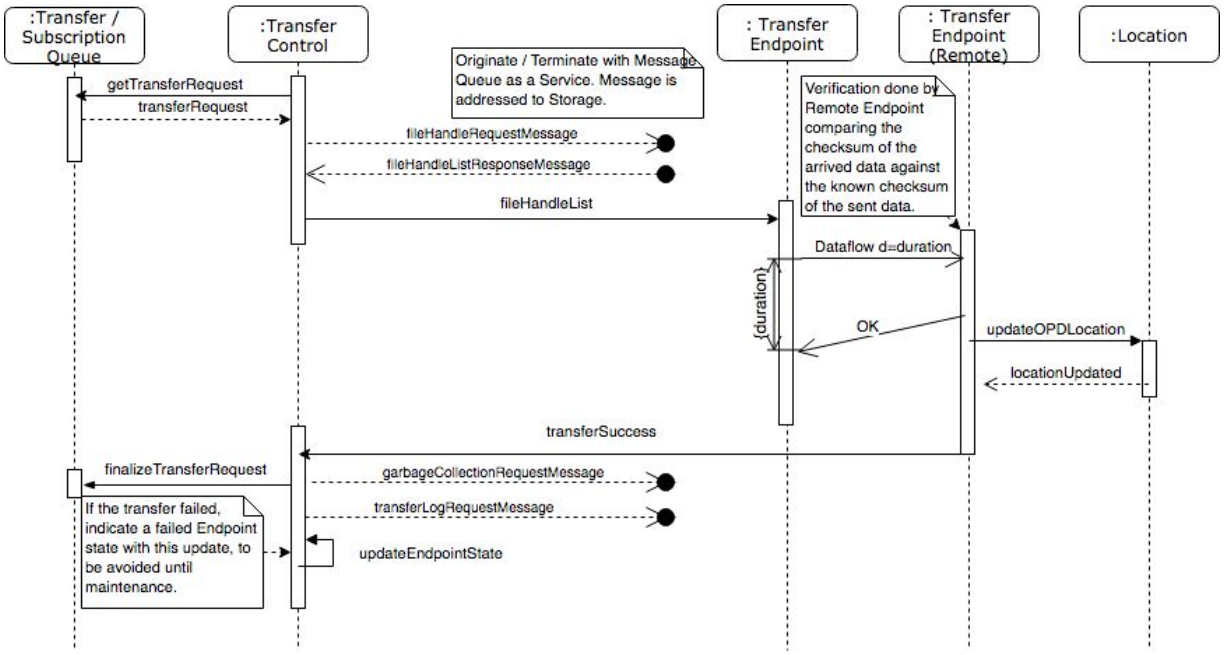

Figure 6: Sequence of actions taken when a transfer request is removed from the data transfer queue.

Transfer Control gets the names of requested data products from the Transfer / Subscription Queue and requests these via the coordination interface. File handles are returned to Transfer Control which then passes these to the Transfer Endpoint to perform the data transfer. The remote Transfer Endpoint updates the Location database once it has received the product and it is ready to share it with other sites. Transfer Control informs the Storage component that it no longer needs access to the file at this time. The status of the transfer is the communicated back to the Transfer / Subscription queue to either remove the request, or mark it as failed. Failed requests could be retried periodically, but that is not shown in Figure 6.

3. Context Diagram

The Context Diagram below depicts the Delivery C&C scope. See also the Operational System C&C View primary representation [AD02].

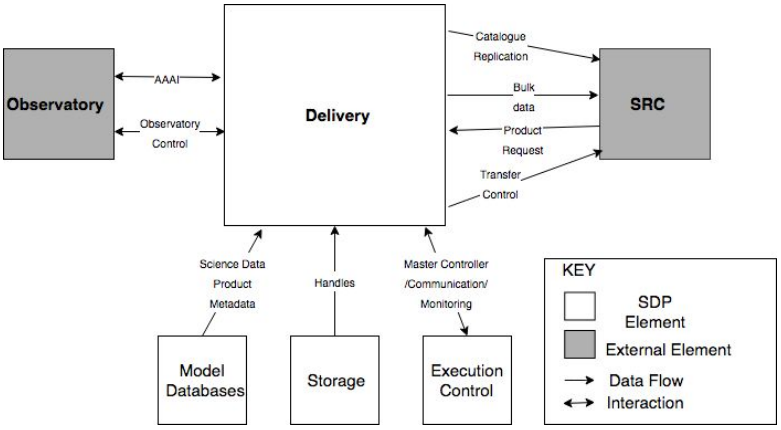

Figure 7: Context Diagram of the Delivery System

4. Variability Guide

The same components will be used at both SDP sites. SRCs need to support the same file transfer protocols and support the location service updates. They also either need to have a set of IVOA services that support the schema used in the Science Data Product Catalogue, or convert the replicated Science Data Product Catalogue to a data model that works with their IVOA service implementation. All sites (SRCs and any other) providing support for science data processing need to provide a common set of services.

5. Rationale

5.1 Drivers

It is undesirable to have end users accessing SDP resources directly, since they are being sized and provisioned just to support the creation of the SDP Data Products and the networks are being sized for transporting these products to SRCs. Therefore, support is given according to the requirements for telescope operators and we have provided support for data transfer requests and monitoring to SRC operators, but not attempted to expose a full set of science services to the the science community as a whole.

Given that the SRC design has been outside of the scope of SDP, and is being done on a longer timescale, we needed a data transport system that can interface with as yet unknown storage systems on remote sites. Therefore we have used a simple data transfer endpoint paradigm that can can be integrated with multiple different local storage systems.

The use of the endpoint paradigm also means that a range of tools including GridFTP service can be deployed, that are capable of the high performance WAN transfers needed to make best use of international research networks.

Given that searches of the Science Data Product Catalogue can closely couple interaction with the tools and services that perform these, replication of the catalogue to sites closer to where these services will run is preferable. Therefore we have catalogue replication as a part of the architecture. We prefer solutions that can be implemented using Open Source software to avoid being locked into proprietary solutions, and have already prototyped the key parts of this system using Open Source software.

We also recognise that future development of the Delivery System will likely be performed in collaboration with CERN. We have used tools developed by the CERN community where possible in our prototyping and this has influenced the architecture.

We want to enable tools and operators at SRCs to locate SDP Data Products that have been replicated to other SRCs, thus enabling them to be accessed without a new request to an SDP site. This is to limit the use of the WAN networks leading out of the Science Processing Centre to mainly handling new product distribution.

The Science Data Product Catalogue has to support IVOA services due to the following level 2 requirement:

> SDP_REQ-710 Science Data Product Catalogue IVOA support. The SDP science catalogue shall contain sufficient meta-data to support standard IVOA queries as a minimum.

Therefore we have contributed what is required for this to the SDP Data Model. Also, the SDP has to support a set of IVOA services to SKA staff due to this L1 requirement:

> SKA1-SYS_REQ-2353 Virtual Observatory interface. The Science Data Products being long-term preserved by the SKA1_Low and SKA1_Mid telescopes shall be accessible via a set of International Virtual Observatory Alliance (IVOA) services and data models. The set includes SIA, TAP, SSA, DataLink and SODA services and the ObsCore data model. Access to SKA1_Low and SKA1_Mid Science Data Products shall only be possible for SKA staff users.

The architecture allows for this through a set of services provided in the Delivery Access component.

5.2 Quality Attribute Scenarios

Scenario/Driver	**Applicability**
SDP_REQ-782 - Power Interruption Survivability	Delivery element able to recognize and recover from an interrupted data transfer.

SDP_REQ-793 - SKA1 Software Quality	Includes copyright notice, software license, and practices that cover work management, code management, documentation, testing, deployment, logging, and alarming as described in [AD01]. The standard manage practices will limit the choices for startup, shutdown, and recovery.

5.2.1 SDP_REQ-285 Accessibility

The SDP shall enable per user access to SDP resources (hardware and software) using the Authentication and Authorisation facilities provided by the SKA.

Scenario Refinement for Scenario 1		
Scenario(s)		All software security issues discovered must be communicated to deployment sites within one day and fixes provided within one week.
Business Goals		Security of SKA and SRC data and services.
Relevant Quality Attributes		Security
Scenario Components	Stimulus	Identified security risk
	Stimulus Source	Security issue discovered from an incident or from code review or notification from external source.
	Environment	Internet facing services, firewalls and intrusion detection systems. Communication should occur during all operating environments.
	Artifact (If Known)	Maintainers of the delivery software
	Response	Maintainers to send out notification to a security communication list and then a software update. Communication sent out regarding fix or mitigation. Software fixes deployed at the SDP sites.
	Response Measure	Communication within one day. Fixes provided within one week. If that is not feasible then provide mitigation within one week until the fix is ready. Communication about fix or mitigation within one week or as soon as available.
Questions		
Issues		

Scenario Refinement for Scenario 1.1		
Scenario(s)		Delivery system software must be secure and protected using SKA authorization.
Business Goals		Enforce data access policy for data products
Relevant Quality Attributes		Security
Scenario Comp	Stimulus	Detection of unauthorised access to data products, intentional or unintentional.
	Stimulus Source	Human or system
	Environment	During normal operation and during maintenance, down-time, etc. Operation with the SKA data access policy.

one nts	Artifact (If Known)	All software that can access data products. Access logs. Data product storage systems (preservation system, delivery system, SRC storage, backup storage, etc.). SKA A&A system.
	Response	Prevent further unauthorised access to data products. Determine the intent of unauthorised access. Record all access to data products. Fix the bug in the system. Inform the SKA Observatory.
	Response Measure	Once unauthorised access has occurred, prevent further unauthorised access within 1 minute until issue is resolved. Communication within one day. Fixes provided within one week. If that is not feasible then provide mitigation within one week until the fix is ready. Communication about fix or mitigation within one week or as soon as available.
Questions		
Issues		

6. Related Views

This view is a decomposition of the Operational System Component and Connector View [AD02]. The System-level Security [RD04] view further describes security issues as relating to the Delivery System. SKA Regional Centre relationship is described in the SKA Regional Centre Requirements [RD02].

7. References

7.1 Applicable Documents

The following documents are applicable to the extent stated herein. In the event of conflict between the contents of the applicable documents and this document, **the applicable documents** shall take precedence.

[AD01] SKA-TEL-SKO-0000661 - Fundamental SKA Software and Hardware Description Language Standards

[AD02] SKA-TEL-SDP-0000013 Operational System Component and Connector View

[AD03] http://www.ivoa.net/documents/TAP/

[AD04] http://www.ivoa.net/documents/latest/SIA.html

[AD05] http://www.ivoa.net/documents/DataLink/

[AD06] http://www.ivoa.net/documents/SSA/

[AD07] SKA-TEL-SDP-0000013 Rev 05 SDP System-level Data Model View

7.2 Reference Documents

The following documents are referenced in this document. In the event of conflict between the contents of the referenced documents and this document, **this document** shall take precedence.

[RD01] gridFTP: http://toolkit.globus.org/toolkit/docs/latest-stable/gridftp/

[RD02] SKA-TEL-SKO-0000735, SKA Regional Centre Requirements, R. C. Bolton and the SRCCG

[RD03] SKA-TEL-SDP-0000038, SDP System Sizing, R. Bolton, F. Graser, F. Malan, R. Nijboer

[RD04] SKA-TEL-SDP-0000013. System-level Security View. Rev 05.

8. Version History

Version	Date of Issue	Prepared by	Comments
04	2017-12-15	R. Simmonds, S. Goliath	Prepared for the SDP M19 deliverable. Supersedes SKA-TEL-SDP-0000025 rev 02.
05	2018-04-23	R. Simmonds, S. Goliath, P. Wortmann, K. Kirkham, S. Sanchez	Prepared for the SDP M20 pre-CDR deliverable.