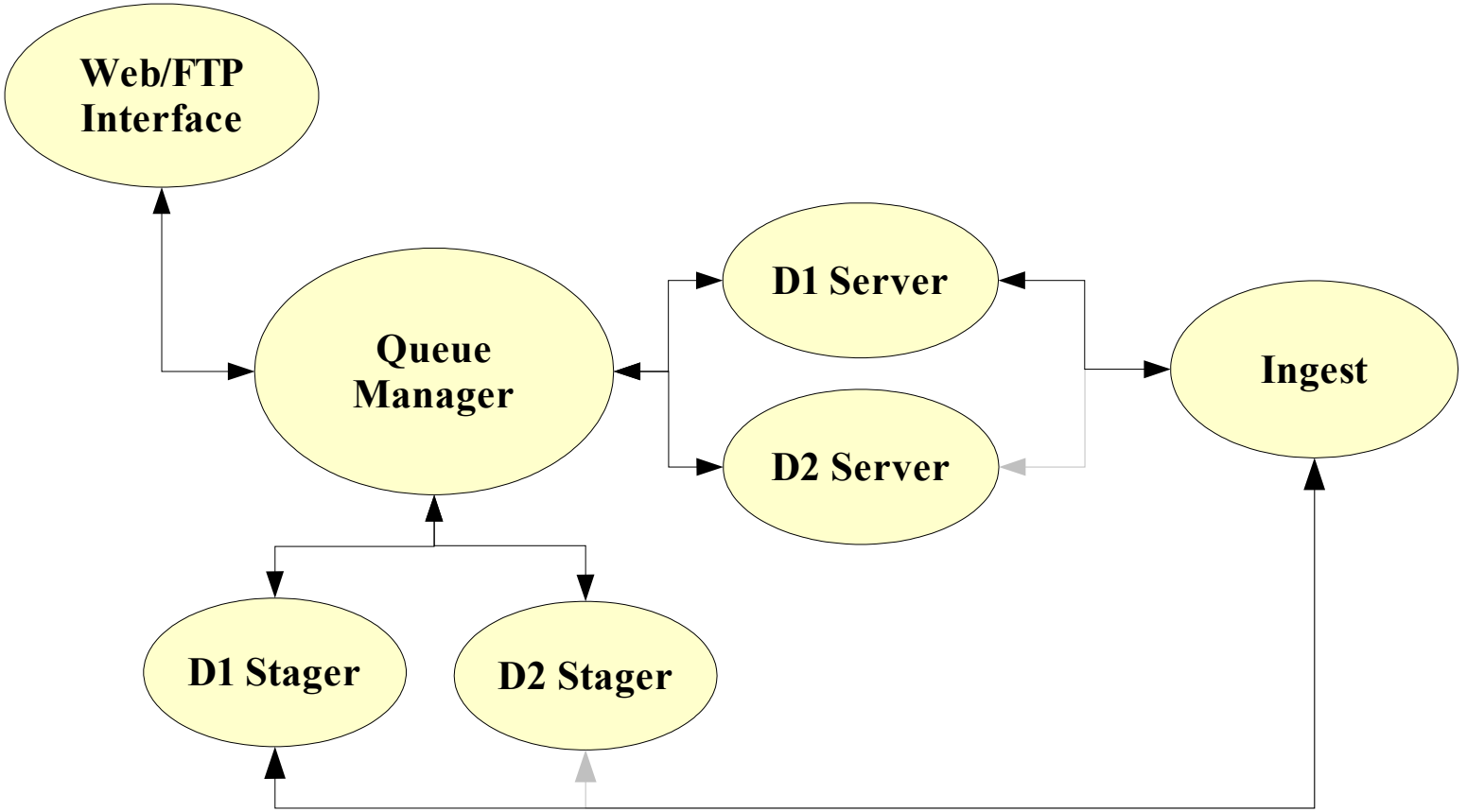
Database Components

Tom Stephens
GSFC

November 2003

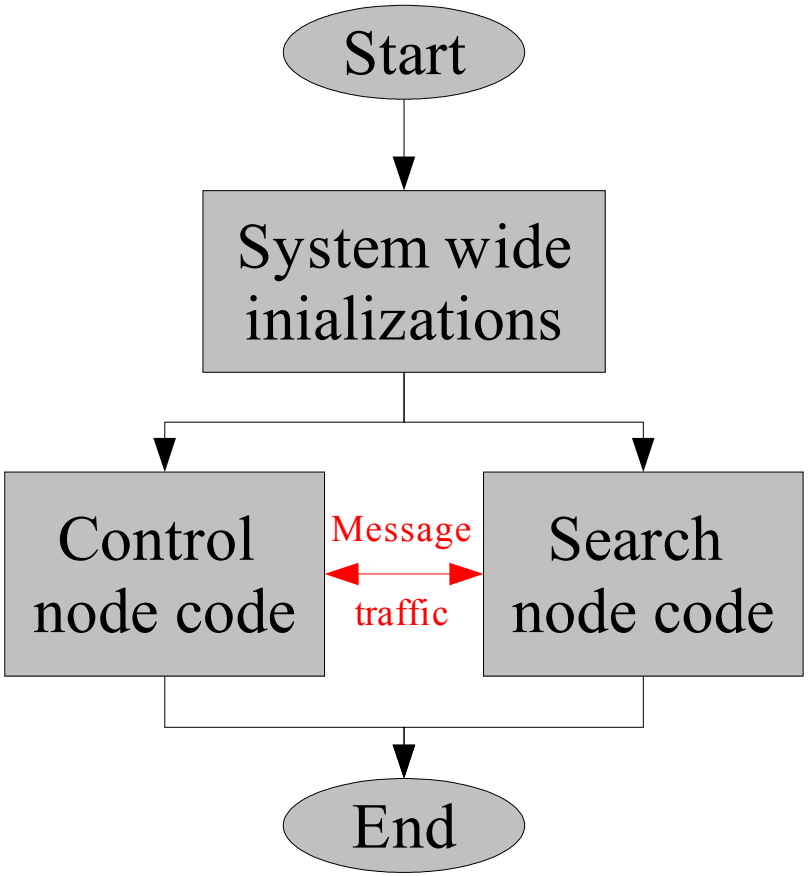
Data Flow


Web/FTP Interface
Queue Manager
D1 Server
D2 Server
Ingest
D1 Stager
D2 Stager

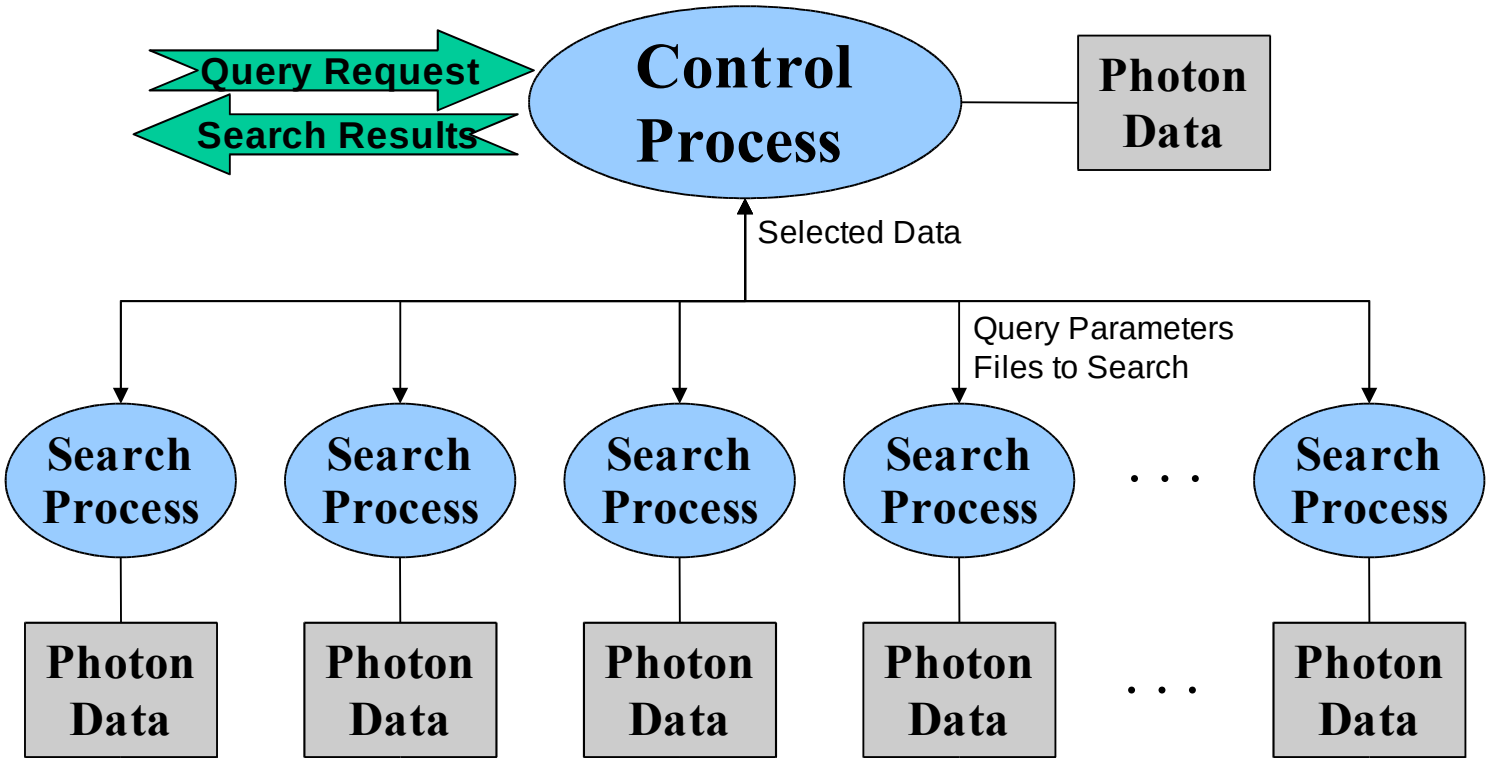
Server Design

- Operates as TCP/IP server for connections from Queue Manager and Ingest program.
- Parallel program using MPICH as message passing interface.
- Scalable to any number of processes, but can run on a single computer.
- Written in C
  - ease of implementation and interface with cfitsio and MPICH libraries
  - Maximum performance

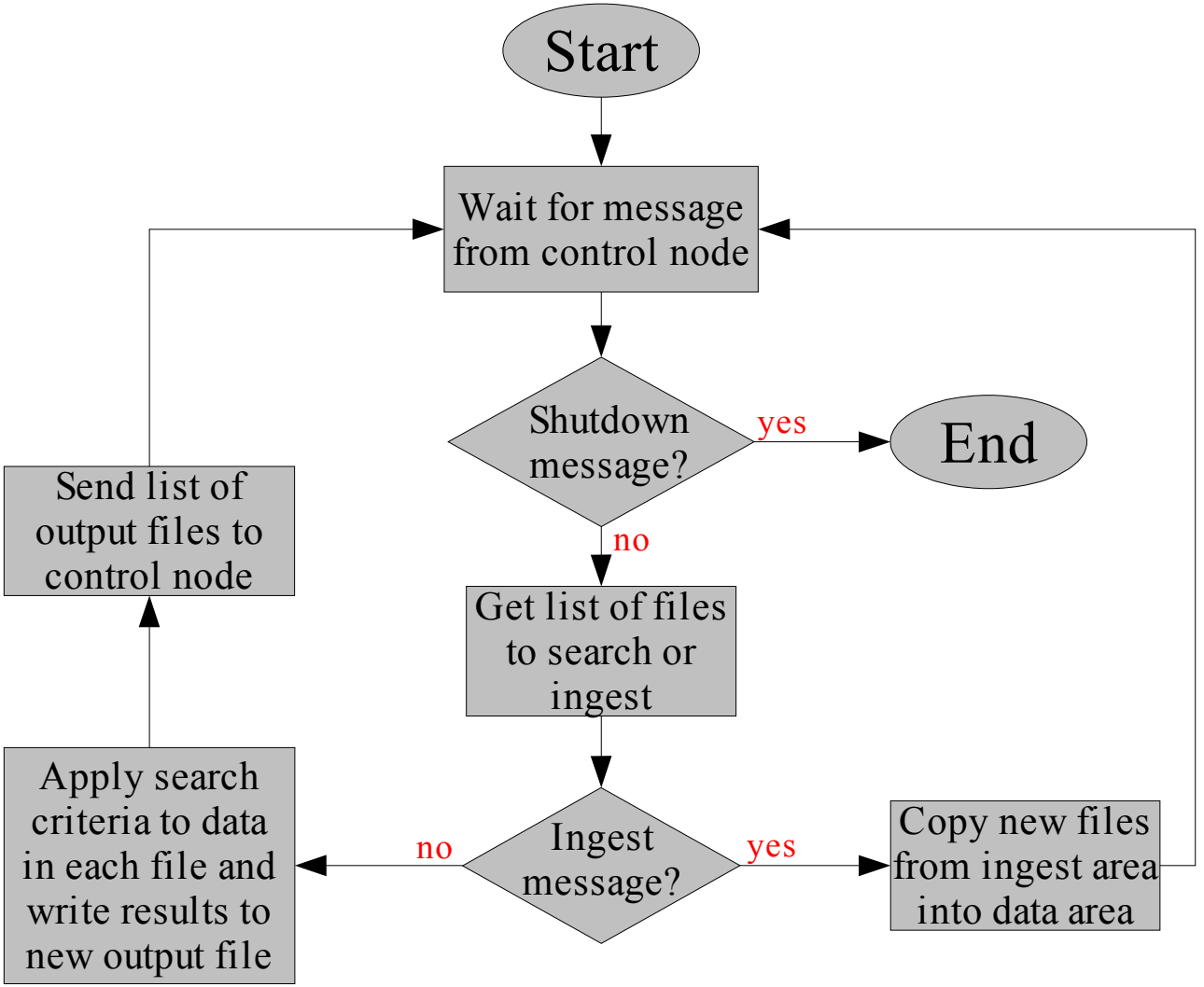
Program Flow I



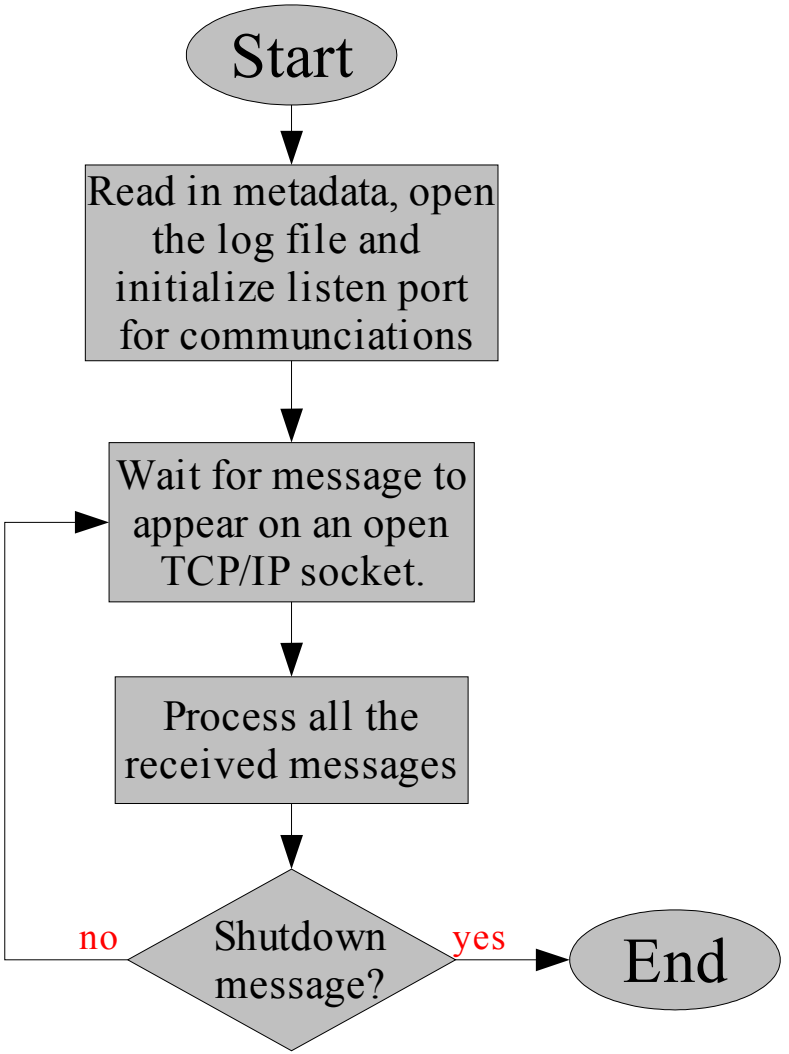
D1/D2 Internal Process Structure



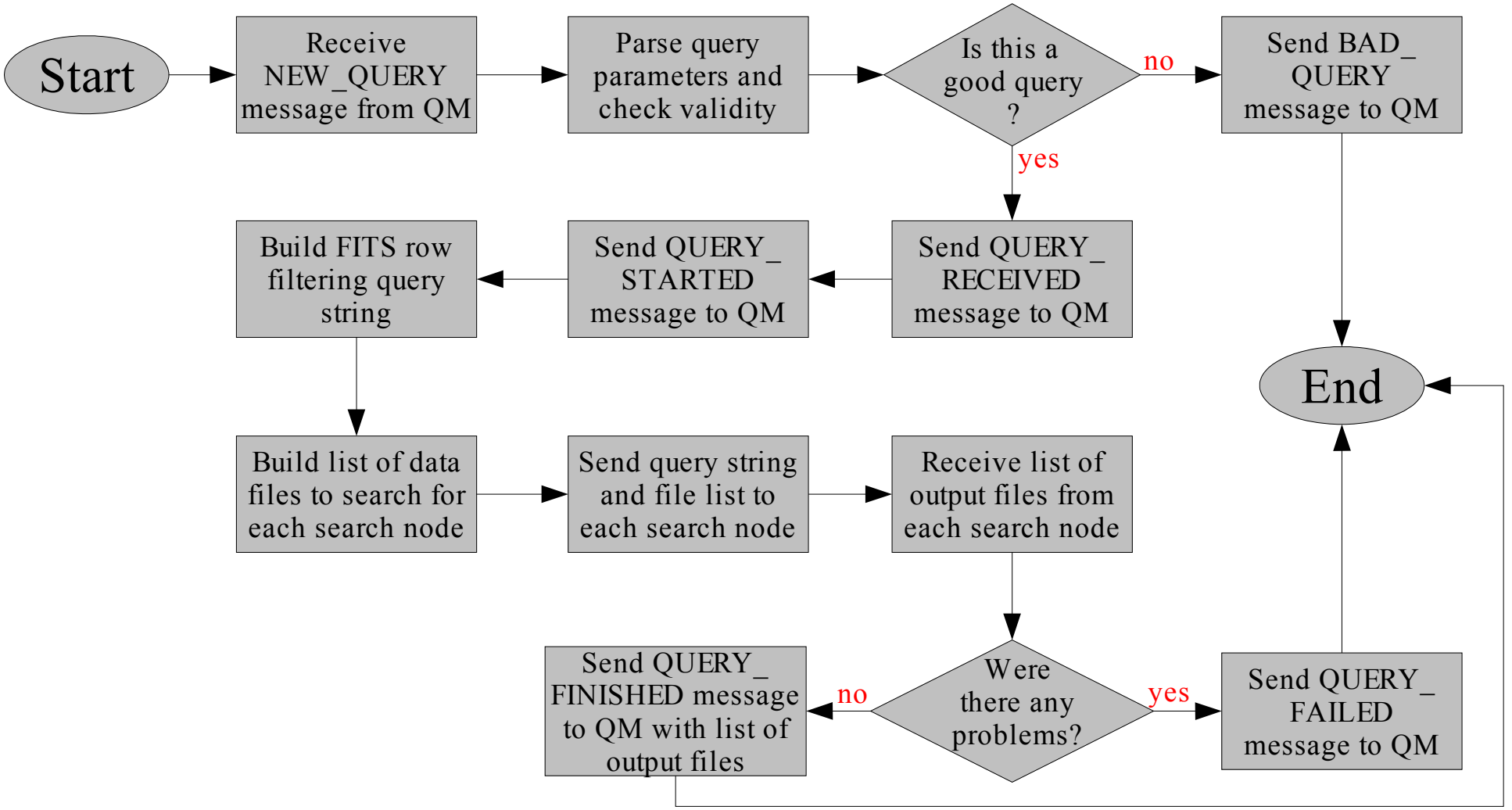
Program Flow II – Search Node



Program Flow III – Control Node



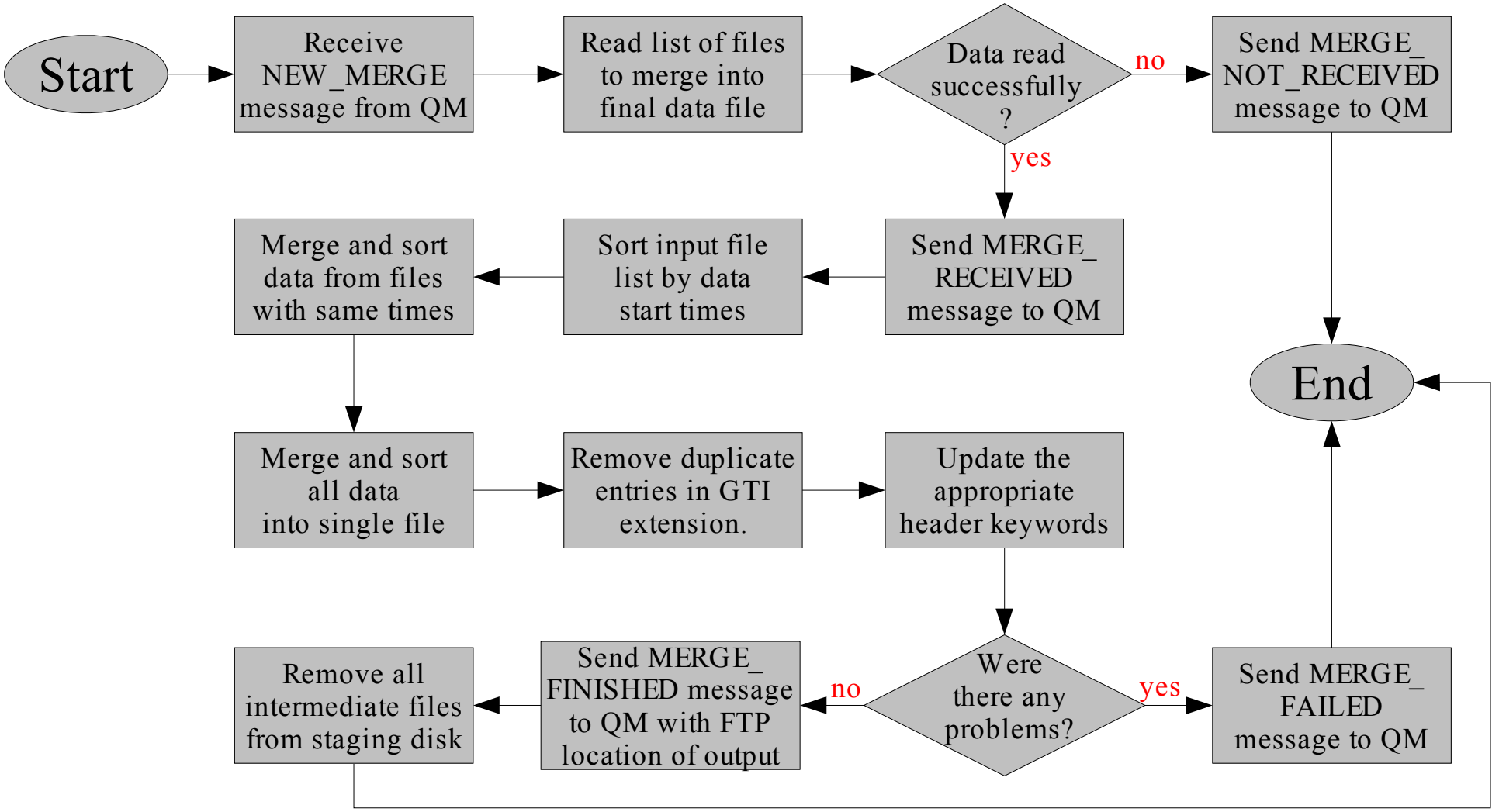
Program Flow IV – Conducting a Search



Stager Design

- Operates as TCP/IP server for connections from Queue Manager and Ingest program.
- Must merge and sort data from search nodes and return an FT1/FT2 file to the user.
- Overall structure similar to control node of server but without parallel processing.

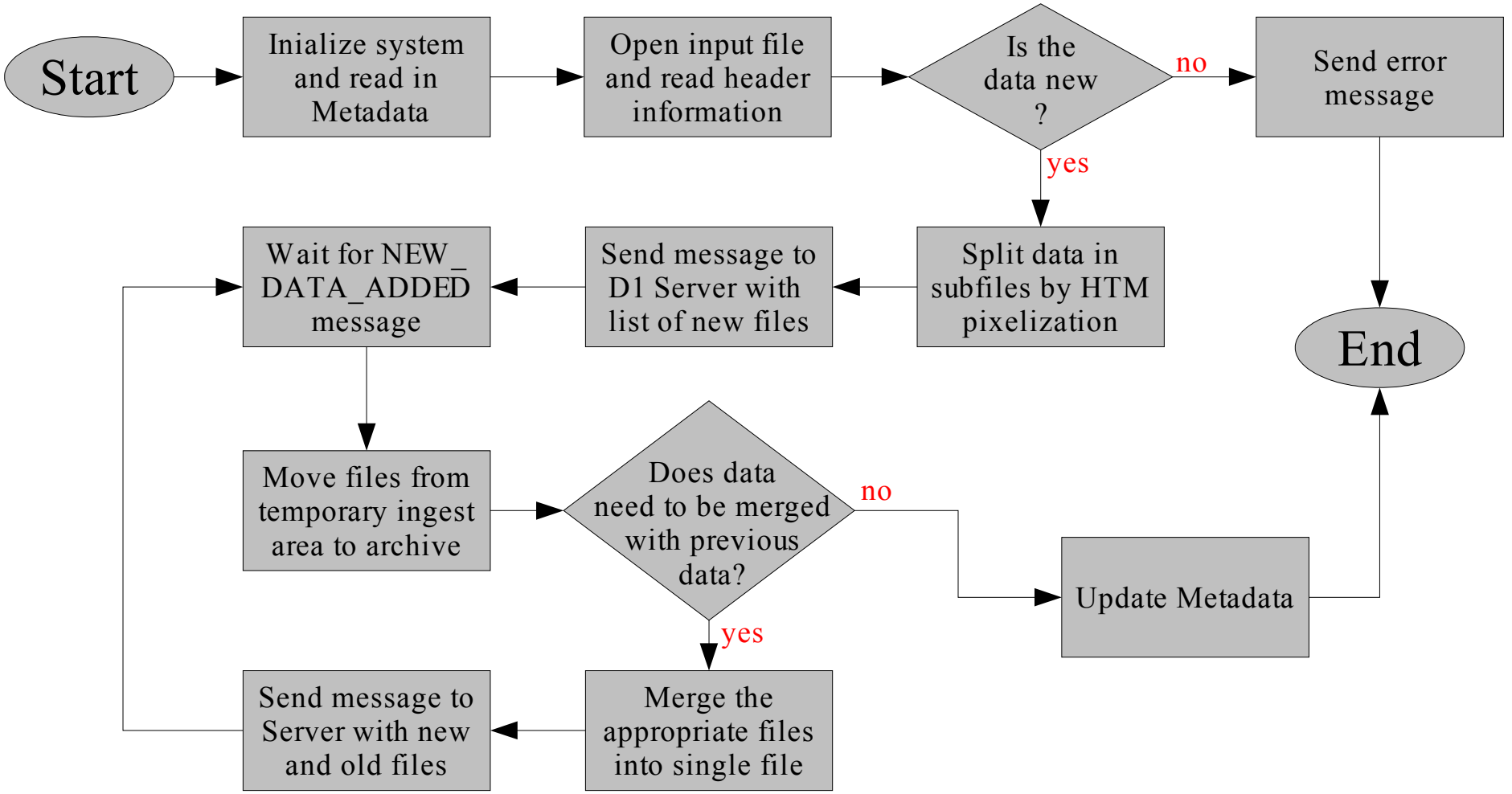
Program Flow – Merging Query Data



D1 Ingest Design

- Operates as TCP/IP client to Server and Stager
- Prepares data output from LAT Level 1 processing pipeline for ingest into databases.
- Breaks sky into regions using Hierarchical Triangular Mesh (HTM) indexing.
- Will probably be rewritten in C++ using Goodi once final decision on HTM is made.

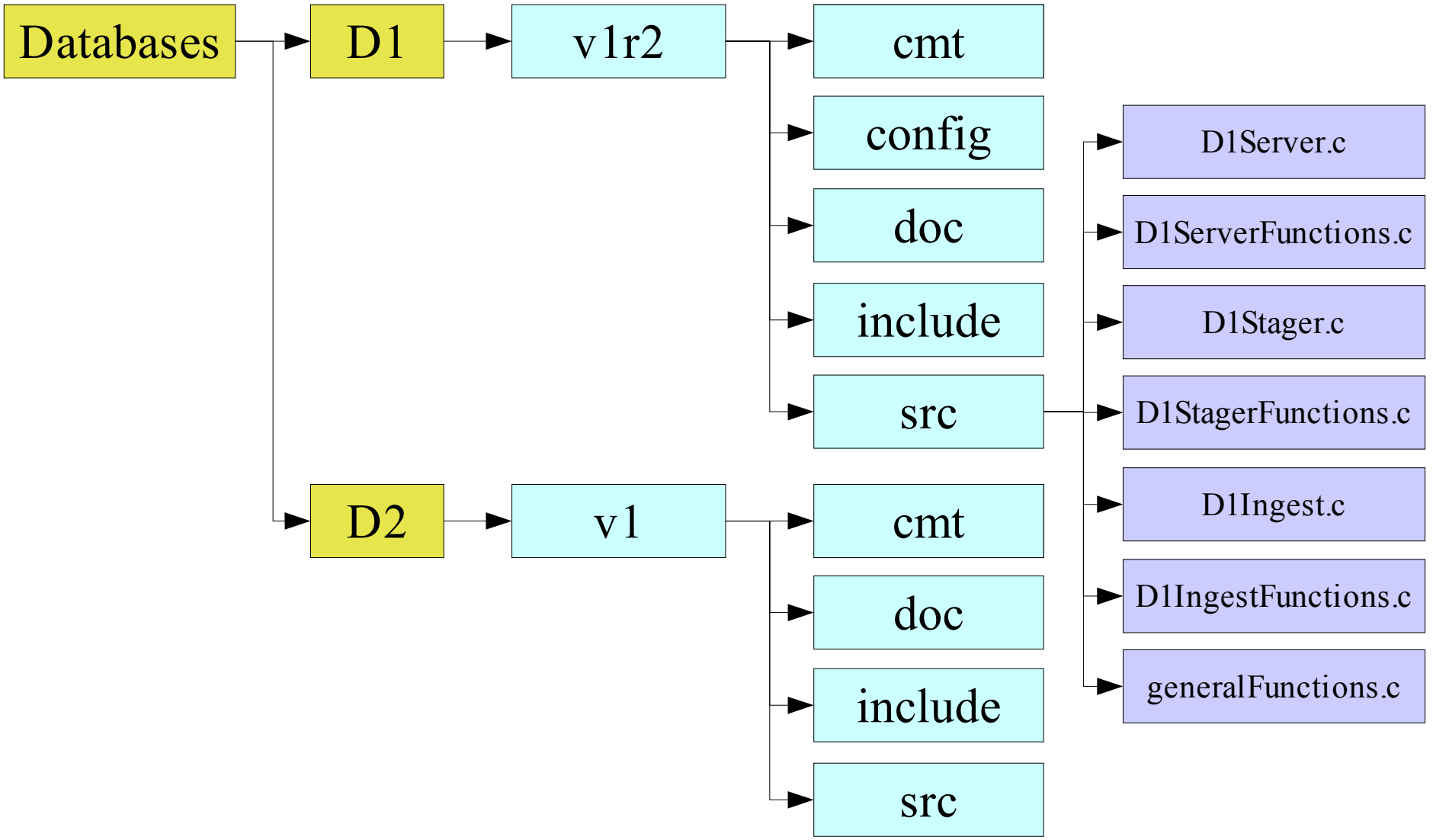
Program Flow – Ingest New Data



Sample Configuration File

```
############################
#  Sample configuration file
#  These are the settings used on the hygd cluster.
#  Any line that begins with a '#' character is a comment and is ignored.
#  Edit the second column to reflect your specific site configuration.
#  The order of the keywords does not matter but spelling does.
############################
HOME_DIR      /data/nodes/GLAST/D1
FINAL_DIR     /data/head/GLAST/repository/D1
FTP_HOST      heasarcdev.gsfc.nasa.gov
URL_PREFIX    ftp://legacy.gsfc.nasa.gov/glast
FTP_DIR       /FTP/glast   #ftp directory on heasarcdev
INGEST_DIR    /data/head/GLAST/ingest/D1
NEW_DATA_DIR  /data/head/GLAST/new/D1
LOG_DIR       /data/head/GLAST/logs
ARCHIVE_DIR   /data/head/GLAST/archive/D1
HEAD_DATA_DIR /data/head/GLAST/D1
SERVER_PORT   45278
STAGER_PORT   45280
```

Code Layout


Databases
D1
v1r2
cmt
config
doc
include
src
D1Server.c
D1ServerFunctions.c
D1Stager.c
D1StagerFunctions.c
D1Ingest.c
D1IngestFunctions.c
generalFunctions.c
D2
v1
cmt
doc
include
src

To Do

- Write ingest for D2
- Add code to ingest reprocessed data
- Refine ingest process for D1
- Benchmark D1 both with and without HTM pixelization