Canada-France-Hawaii Telescope Corporation
65–1238 Mamalahoa Hwy. Kamuela, Hawaii 96743 USA

Société du Télescope Canada-France-Hawaii
Telephone (808) 885-7944 FAX (808) 885-7288

CFHTLS operations at CFHT – Section 1.2.1 of SAC's CFHTLS mid-term review

To: CFHT's Scientific Advisory Committee (SAC)
From: Jean-Charles Cuillandre, Christian Veillet, Pierre Martin
Date: March 15, 2005

Abstract & Content

This document presents how the CFHT Legacy Survey (CFHTLS) is operated at CFHT under the New Observing Process. CFHT's responsibility in the CFHTLS joined effort with the CADC and Terapix to provide data to the Canadian and French communities (and world later on) is to gather the raw data with the MegaPrime/MegaCam instrument under Queued Service Observations mode, to calibrate and remove the instrumental signature from the data with the Elixir pipeline, to collect ancillary data and ship them along the FITS Elixir data (what is dumbed as a whole as the CFHT data products), through the CFHT Data Archiving and Distribution Service to the Canadian Archiving Data Center. After describing the scope of each of these three components and how they have evolved under the impulse of the CFHTLS, this document focuses on identifying the reasons why over the past three semesters since the survey started, the CFHTLS data have not been gathered as fast as initially expected. The main culprit appears to be the Mauna Kea weather, but the operationa and intrumental instrument overheads also play a role. Only by decreasing the observing overheads can CFHT get close (5.5 to 6.0 hours per night) to the number of hours of validated data per night that was advertised at the time the CFHTLS was being defined (6.5 hours per night).

1 How it became possible: the New Observing Process (NOP)

In an age where most major ground based astronomical facilities turned their focus on service observing and queued prioritized observations, CFHT decided to produce a large effort in 1999 to put such a scheme in place for an instrument monopolizing most of the telescope time throughout the year (40%): the CFH12K, a 100 million pixels camera with a field of view of 42 by 28 square arcminutes. Wide-field high resolution imaging had then already been identified as the best niche for CFHT to keep its leading scientific position in the era of the very large telescopes. At the time of the first light of the CFH12K in January 1999, the design of its follow-up was already well advanced: the MegaCam camera (funded and built by the CEA-DAPNIA, France) to be housed in a brand new top-end, MegaPrime (funded by CNRC and CNRS, built by CFHT, HIA and OPM). MegaPrime, with its large wide-field corrector built by SAGEM was to provide an increased field of view and better image quality compared to the old prime focus, up to a radius of 0.7 degree from the center of the mosaic (MegaCam's field of view is 1 square degree). Note here that a condition set by CEA-DAPNIA to fund and build MegaCam was that a major survey would have to be offered to the CFHT community, this was the seed of what would become the CFHT Legacy Survey.

It was clear that in order to optimize the scientific productivity of these major investments (MegaPrime being the largest and most expensive instrument ever built for CFHT was also expected to use as much as 60–65% of the telescope time), the service observing scheme had to be put in place, operated and smoothed out before MegaPrime could start its operations. Soon after the CFH12K had started its operation on the telescope in 1999, the brainstorming started at CFHT to set up the new observing scheme. It was decided of not just taking data in service mode, but to look at the process of gathering data from CFHT as a whole: from the submission of a time proposal up to the processing and delivery of the data to the scientists at their home institutions. This was the birth of the CFHT's New Observing Process (NOP) which is made of four main components:

- Queued Service Observations (**QSO**)
- **Elixir**, the data processing and calibration pipeline
- Data Archiving and Distribution Service (**DADS**)
- New Environment for Observing (**NEO**)

The NOP was put in operation in January 2001, and operated on CFH12K for two years before MegaPrime became the new official imager in February 2003. The following sections summarize the main characteristics of the four components as they run today for MegaPrime, with more attention given to the elements pertaining to the operation of the CFHTLS (to be discussed in a later section).

1.1 MegaPrime/MegaCam

MegaPrime is the newest CFHT wide-field imager. It is the composed of a new prime focus upper end meant to provide higher image quality to the camera it houses: MegaCam. This camera is made of 36 2048×4612 pixel CCDs, for a total of 340 Megapixels. The camera samples nicely the median seeing at CFHT (0.7 arcsec. in the R band) with 0.187 arcsecond wide pixel. It provides a field of view close to 1 square degree and is sensitive from the near-UV (U band) to the near infrared (Z band). MegaCam uses a set of Sloan filters (u*g'r'i'z'). MegaPrime is also equipped with an image stabilizing unit meant to reduce the effects of wind shake. MegaPrime is operated exclusively through the NOP.

1.2 QSO

The QSO system consists of a suite of tools and softwares operated by a dedicated group of astronomers and service observers at CFHT. The system functions range from the submission of observing programs, the preparation of queues on a per night basis, the execution of the queues throughout the night, and finally the evaluation of the observations.

A team of four resident astronomers at CFHT share the duty of preparing the queues and validating the observations (the Queue Coordinators). This requires several hours of attentive work every day, and the duty rotates every 4 to 5 days. A team of two dedicated service observers (the Queue Observers) with occasional backup from the pool of observing assistants, aka the telescope operators, run the queues at night, jumping from one queue to another depending on the observing conditions. The observing conditions are defined as:

- Seeing (image quality measured at the center of the field)
- Sky background (average level over several CCDs)
- Sky transparency (a tool called SkyProbe tracks this within 2%)
- Time constraints (follow-up programs, e.g. KBOs or SNe)

About five different queues covering different sky conditions are prepared every day to optimize telescope time and also to minimize overheads as much as possible (number of filter changes, number of areas of sky visited). This allows for the most adequate programs suited for the given observing conditions to be observed. There are however overheads associated with jumping from one queue to another, hence that option is also part of the equation to optimize the use of the telescope time. After the images have been obtained the first three parameters listed above are measured to allow later on the QSO coordinator validating or disqualifying that given exposure. If validated, the time is charged to the program (integration time plus the overhead of 40 seconds for MegaCam). The conditions remain usually stable during a set of exposures taken for a given program, but sometimes conditions evolve and force several exposures to be dismissed. This is what defines the Queue validation efficiency which sits typically for MegaCam at 80 to 85% - the best ever achieved was 92% for CFH12K, 100% is impossible as conditions are always bound to evolve erratically sometimes during some nights.

During the time QSO was operated on CFH12K (2001-2002), several time constrained programs had been successfully executed (e.g. satellites of Jupiter, KBOs, Supernovae), proving then that only service observing as provided by QSO could carry on successfully two of the three main programs identified for the CFHT Legacy Survey: the SNLS and the Very Wide component.

Following the SAC and BoD recommendations, the main goal of QSO is to achieve a very good completion level on grade A programs. QSO is able to achieve this regularly with A programs completed at a level higher than 90% when the weather is within reasonable statistics. A programs represent approximately 35 to 40% of the queue time for a given agency, and B programs fill the other 65 to 60%. The number of nights programmed on the telescope is directly derived from the total amount of hours from A and B programs requested for all agencies: total number of nights divided by 6.0 hours of validated hours per night (versus 6.5 on the first semesters, more on this in section 3).

A highly important scheduling constraint for the QSO coordinators at the granularity of a QSO observing run (which lasts typically 17 to 20 nights, but some have been as short as 10 days as the MOS instrument requires dark time also), is to balance the agencies: Canada, France, Hawaii, Korea, Taiwan, and now the CFHTLS ”agency”. Queues will be tailored throughout a run primarily to keep the balance of the observed time matching the balance of the requested time for each agency (typically, the CFHTLS represents 50% of the whole QSO allocation). This ensures all agencies suffer equally from the bad weather and technical problems.

The success of the QSO operation is unquestionable: over the past four years it has proven that the best of the telescope and observing conditions can be used to gather data based on the merit set by the Time Allocation Committees (TACs) ranking. Also, the agency balancing is perfectly respected when comes the end of a semester. Section 3 presents the challenges QSO faces with the execution of the CFHTLS, and the realities of observing where bad weather (mostly) and lack of efficiency of the observing chain hamper the gathering of data.

1.3 Elixir

At the time the first wide-field imager saw first light at CFHT, MOCAM (16 Mpx) in 1994, little was known about the subtle effects involved in using mosaics of detectors (observing strategies, data processing), or pushing down the limits on the astrometry and photometry on large field of views. When the UH8K (64 Mpx) became available at CFHT, standard image processing software proved to be inadequate and new tools had to be developed to allow fast detrending of the data (pre-processing). When CFH12K arrived with its 100 million pixels (200 Mbytes files), it was clear that a large fraction of the CFHT community simply could not handle the data the old way: starting from raw frames. When the NOP project was started with the idea of providing fully processed data (removal of the instrumental signature) to the user, the scope of the data processing was extended to a full per CCD astrometric and photometric calibration of the data (no stacking and global astrometric calibration was envisioned as the Terapix data center was focusing major efforts on this specific step). The suite of softwares composing the entire pipeline is called "Elixir".

From the Elixir viewpoint, the advantage of the QSO operations is that all data are taken with equally experienced observers and the data gathered throughout a run are a lot more uniform than what was achieved by a set of occasional visitor observers. This allows for high quality master flat-fields and fringe frames to be created per observing run. Also, photometric standards are only observed when conditions are photometric, allowing for a zero point to be derived per run with several reliable measurements.

CFHT committed to provide images fully processed (full removal of the instrumental signature), flat photometrically to within 1%, and astrometrically calibrated to within 0.2 arcsecond. The CFH12K had little fringing and the recipes developed at that time showed their limits on the MegaCam data which have fairly strong fringing in the i' (6%) and z' (15%) bands. A fancier correction algorithm is under development. For the flatness of the photometry (that is the flux from a given object should be the same on any location of the field of view), Elixir uses a "photometric grid" allowing the calibration of the flat-field scattered light effects, the change of optical scale from center to edge, and the apparent broadening of the filter bandwidth from center to edge. Such map is convolved into the master twilight flat-field, allowing all effects to be corrected in a single step when the data is pre-processed. Elixir's astrometry runs exclusively on a per CCD basis and it can happen that given CCDs (specially in the u* band) don't have enough stars to allow Elixir to match the USNO B1 catalog (in which case the initial World Coordinate System derived from the telescope pointing remains unchanged), or worse, create a wrong astrometric solution for that CCD (a patch has been applied in Feb. 2005 to reduce further this fairly rare case).

Elixir is basically a fully automated process requiring little supervision. During a run, real time statistics are derived from acquired data (seeing, sky level) and fed back to the QSO tools, allowing the observer & coordinator to evaluate the quality of the data. At the end of an observing run, the data are processed following four major steps:

- Create master detrend frames: master bias, darks, twilight flat-fields, fringe frames (i' & z')
- Derive the astrometry of all images on a per CCD basis.
- Derive the zero points for all filters.
- Qualify the science data based on the three previous steps.

For a typical MegaCam run, about 1,700 images (all image types) are obtained, representing more than one terabyte of data. Elixir has to go through each single of them for the various steps listed above and thanks to increased computing facilities at the time MegaCam became operational (namely five dual 2.4 GHz processors PC loaded with memory and disk space), a complete run can be fully processed (all 4 steps) in typically five to six days. There is human quality control after each of these individual steps to ensure all went well. The current support for the Elixir operations at CFHT rests on a single resident astronomer. All Elixir operations can be remotely activated, controlled and checked. After the run processing, all the Elixir data products (image statistics database, photometric database, detrend database) are open and can be used by DADS to proceed with the data distribution.

A by-product of the main Elixir for CFH12K was the development of SkyProbe, a small CCD camera with a large field of view using Hipparcos' Tycho V-band photometry catalog to measure the absolute transmission of the atmosphere where the telescope points. The accuracy is within a few percent and has proven in 99% of the cases to be able to tell if the conditions were photometric or not. This tool has become an essential part of the CFHT nighttime operations, and a second channel in the B band in under construction for backup and cross comparison.

1.4 DADS

The role of the DADS within the NOP is to archive the raw data, send the raw data to the Canadian Archiving Data Center (CADC), send the Elixir processed data on DLT tapes to the Principal Investigators at their home institution (to CADC for the CFHTLS Elixir data).

Setting up the NOP at CFHT was also about increasing the value of the FITS data by accompanying them with a set of ancillary information. This data set is composed of weather statistics (including SkyProbe data for example), observing logs, comments from the QSO team, and Elixir statistics. All this info is integrated within a HTML template distributed on a CDROM to the PI for easy browsing (all this has been ported also to FITS tables for archiving at CADC).

DADS gets going with the archiving process as soon as a new image is obtained at the telescope: the image is copied automatically to a different host at the summit, compressed and transfered to the disks in Waimea through a T1 line. At that point data are immediately saved on two individual Super DLT tapes, read back and checked. One of the set of DLTs is then shipped via FedEx to CADC every 5 to 6 days.

For the distribution process, DADS has to wait for the green light from Elixir at the end of an observing run if the PI has asked for immediate access to the processed data. Otherwise data are processed in a bulk at the end of the semester and shipped to individual PIs.

The DADS system in Waimea consists currently of 15 Terabytes of disk storage distributed on 15 nodes. DADS uses Elixir's parallel processing capabilities only at the time the data need to be shipped (i.e. only science raw data are kept on disk at all time). It takes typically 3 minutes for Elixir to go from the raw file to the final processed image with all the updated astrometry and photometry keywords.

As for all the other NOP components, DADS does not have a full time dedicated manpower but two engineers from the software group who overview the operations.

1.5 NEO

NEO is primarily an interface between QSO (the observer) and the instrument plus telescope. It defines as input a simple command line oriented language, it returns ASCII formated info to the requests and produces FITS files as output. The NEO interfaces were developed in order to minimize the overheads associated with the instrument and telescope control. More generally NEO also covers a suite of observing tools like the exposure time calculator DIET for example. NEO is exclusively a development project and does not have an operational load like QSO, Elixir and DADS. It involves many engineers across several groups at CFHT.

2 The CFHTLS: setting new standards in the NOP chain

The CFHTLS was being defined by the MegaCam Survey Working Group (MSWG, the entity that was to become today's Steering Group) as the NOP successfully operated with the CFH12K. Much was learned

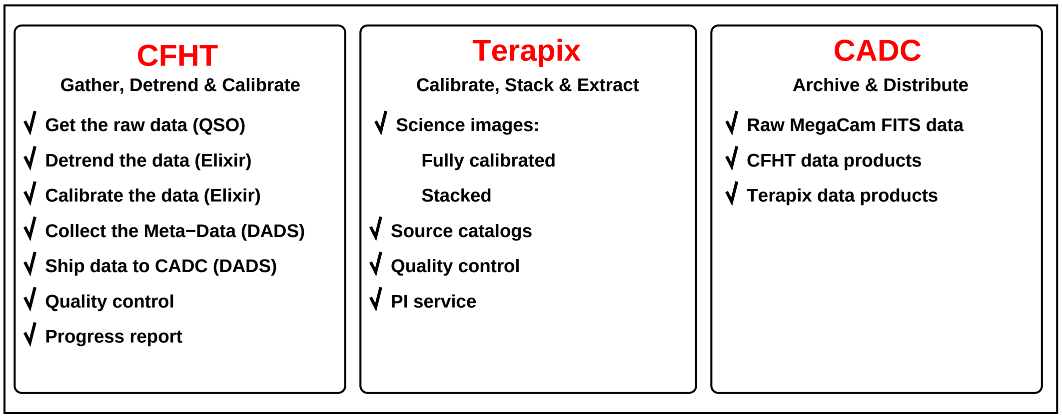

Figure 1: functions of the entities serving the CFHTLS communities

on the feasibility of some scientific programs under QSO, most important of all that time constrained programs could be done successful under such an observing mode. As the CFHTLS became a reality in 2001 after being approved by the SAC and the Board of Directors, there was no real concern about the evolution of the NOP from the CFH12K to MegaCam to support the increasing number of scientific programs, including the CFHTLS, the largest observing program ever conducted at CFHT.

The concerns came from the largest scope CFHT had to approach on its operation: the CFHTLS data flow involved new partners. The CADC had to serve the Canadian and French communities with CFHT data products as quickly as possible after they've been acquired. The Terapix data center was becoming fully part of the chain with the production of stacked calibrated images and catalogs, with CADC still playing its exclusive role of archiving and distributing the Terapix data products.

The following sections follow the CFHTLS Data Flow structure and develop for each step the impact and challenges the CFHTLS has created. As for the evolution of the CFHT data products (Elixir data and ancillary data), it shows how requirements set by the CFHTLS have also served PI programs with higher quality data and services.

2.1 The challenges of the QSO operation in the CFHTLS era

2.1.1 Standard interfacing with the QSO system

Let us start this section with a quick reminder on some QSO functionalities: after PIs have been granted time on MegaPrime by the TACs, the QSO's "Phase 2" (PH2, a web-based application populating a relational database) is open to let them enter the specific details of their observing program:

- Instrument configurations
- Target coordinates
- Observing constraints: seeing, sky background, sky transparency
- Organization in blocks and groups of observations
- Specific comments regarding their program

PH2 is open for a limited amount of time prior the beginning of a new semester (Semester A = February to July, Semester B = August to January).

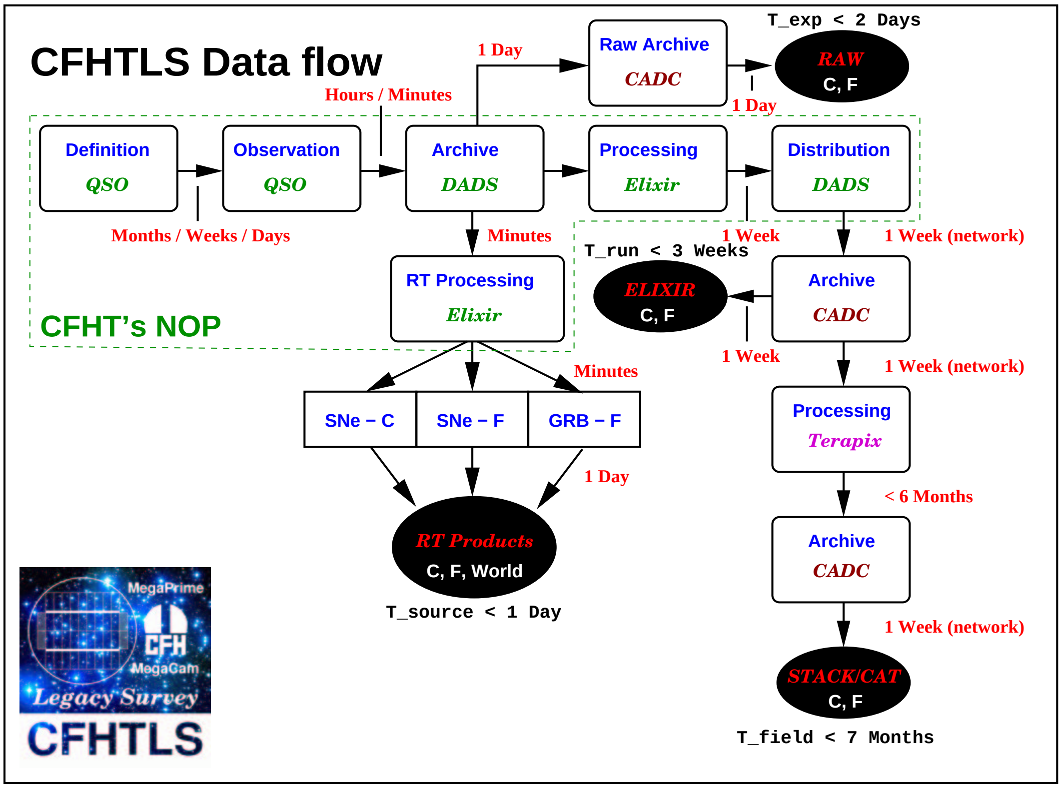

Figure 2: the CFHTLS Data Flow

2.1.2 A QSO privilege to the CFHTLS: dynamic PH2

The complexity of the Legacy Survey time constraint programs however pointed out even before the observations started in 2003, that regular updates of the specifications of the required CFHTLS observations to be executed by QSO were necessary. A granularity of a few days is not convenient as QSO coordinators rely on strategies at the observing run scale. The following approach was adopted over the course of the first semesters of operations: all three components (especially the Very Wide) can update their observations request between two observing runs, CFHT wishing to have this done at least 5 days before the new observing run starts. This is an important privilege, which is also given to certain C-F-H-K-T PIs with time constraint programs.

For the Deep-SNLS survey, the specifications of the observations per field and filter are pretty much the same from run to run but it is the priority within the groups of observations (per field and filter) throughout the run which can quickly evolve, especially when the bad weather alters the observing efficiency of the telescope. A lot of communication via email was taking place initially between the Deep-SNLS coordinator and the QSO team, but this had now been replaced by simple directions updated daily on a web site located in Victoria (home institution of Deep-SNLS coordinator), based on the data obtained (or not in case of bad conditions) the night(s) before (all coordinators have access in a timely manner and in various ways to precise information pertaining to the CFHTLS data acquired during the current observing run – more on this in section 4).

2.1.3 Priorities within the CFHTLS

The previous subsection shows that each component of the survey has the capabilities of changing, altering, and updating its observing strategy from run to run, and during a run. However as the CFHTLS

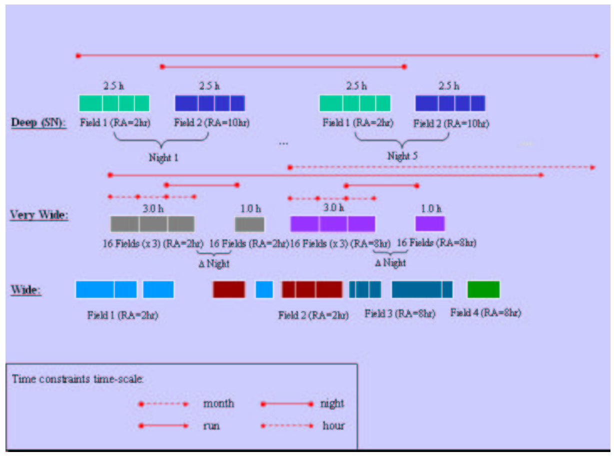

Figure 3: example of possible time clashes between CFHTLS components

program as a whole has the status of an agency, and since the balance of time between agencies is an important constraint (something it always achieves within 1 percent at the end of each semester), it is the CFHTLS coordinators' responsibility, in agreement with the other SG members, to ensure that the balance between the three components of the CFHTLS is respected. The QSO coordinators play of course a crucial role at getting the proper observations obtained but due to the complexity of their task (dealing with many programs from several agencies), the CFHTLS coordinators eventually acknowledged that it was their responsibility in the end to ensure the proper time balancing within the CFHTLS program. It took about a year of trial of various strategies before the current strategy was adopted: a week prior each observing run, the CFHTLS coordinators meet through a conference call and define the CFHTLS strategy for the coming observing run. The result of their discussion is passed on to the QSO team via email in simple terms defining the main guidelines of the CFHTLS agency for the coming observing run. The decisions taken by the CFHTLS coordinators are based on the following criteria:

- Field availability
- Time constraints
- What was obtained thus far (past runs status)
- Status of the instrument (failure, or image quality for example)
- Goals for time balancing between the three survey components
- Goals for time balancing between filters within a given survey

Figure 3 gives an visual example of the competition between the three CFHTLS components within nights of a given observing run (the different colors represent different filters). It is clear there is a competition on the Right Ascension for example on this diagram, a fairly common situation for the CFHTLS (see figure 4 for a distribution of the fields on the sky). Note that figure 3 does not show any PI programs, some of which have time constraints, or are in direct competition withe CFHTLS on the hour angle, equally difficult to fulfil! (see the individual CFHTLS component reports for discussion on the observing strategies)

2.1.4 The sources of the current survey efficiency problem

The CFHTLS is a very complex program highly susceptible to the instrument status, e.g. image quality that forced the Wide survey to start slow up to the end of 2003B. The weather is the most determining factor as bad conditions do not affect all three survey components equally: due to its time constraints, the Deep survey tends to take prominence over the other two components, something that shows clearly in the current global statistics of the survey with the Deep survey ahead of its goal fraction by 13%. The

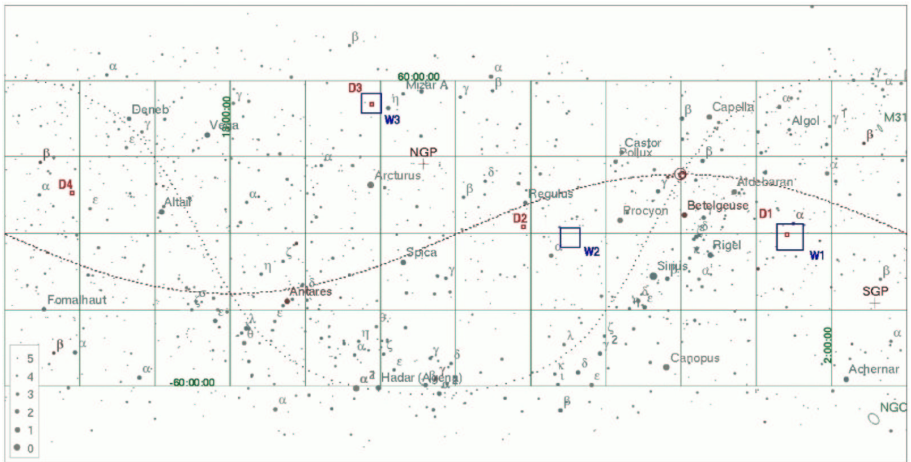

Figure 4: Location of the CFHTLS fields, a source for RA conflicts.

following table presents the status of the CFHTLS observations since the beginning of the survey on May 30th 2003 (the integration time follows the QSO metric with the inclusion of an extra 40 seconds overhead for each exposure). More on these issues can be found in section 3.

Status of the CFHTLS as of Feb. 2005

Survey Component	Deep	Wide	Very Wide
Total integration [validated exp.]	343.8 hr	145.6 hr	109.9 hr
Number of validated exposures	2898	959	2484
Current fraction of CFHTLS	57.4%	24.3%	18.3%
Target fraction of CFHTLS	44.0%	34.0%	22.0%

2.2 DADS

2.2.1 The Interface Control Document

Prior the CFHTLS, the only interface used to exchange data between the CFHT and CADC was FITS, but with the addition of the ancillary data (weather, etc...) it became clear that a clear interface had to be defined between the entities serving the CFHTLS communities. CADC led the effort with the development of the Interface Control Document (ICD), a reference for the format, organization and exchange of data. It was decided that FITS would remain the main interface, with all the ancillary data coded in machine readable FITS tables. This forced a fairly major remodelling of the DADS software that was used to prepare ancillary data for the PIs in human readable format (which is however still preserved for the CDROM distribution).

2.2.2 Summit to Waimea raw data transfer

The Deep-SNLS program calls for a quick identification of the supernovae candidates after the data have been gathered at the telescope. For strategic reasons (support, maintenance), the Real-Time Analysis Systems (RTAS) from the Deep-SNLS teams are located at the CFHT headquarters in Waimea. This means that the raw data have to come down from the summit as fast as possible. DADS setup a scheme in which data from the Deep-SNLS program have priority over all the other MegaCam data for the transfer to Waimea. Later on, a lossless compression scheme was added, resulting in images being archived in Waimea within minutes after having been acquired on the telescope.

2.2.3 Supporting Elixir developments

It takes time and a lot of data mining to derive the best recipes for data processing of a given instrument. It was clear from the beginning of the CFHTLS that the Elixir recipes would improve over time, calling for successive re-release of the data. In order to support the Elixir needs for accessing data over a large timescale, DADS has beefed up its storage capabilities to 15 Terabytes, allowing more than one year of MegaCam data to be available.

2.2.4 Network transfer to CADC

In the past, raw data used to be exclusively sent to CADC on tapes. However, users of data from the Very Wide survey needed access to the data within days, a constraint not strong enough for them to install a RTAS in Waimea. The frequency of the tape shipping from CFHT to CADC was not adequate, hence CFHT setup with CADC a complete scheme for transferring data through the network. CADC being capable of ingesting and publishing an image within minutes of reception, the timescale for having the CFHTLS raw data available to the users at CADC melted down to less than a day after the acquisition on the sky.

2.3 Elixir

2.3.1 The Interface Control Document

As part as the new standards set by the CFHTLS, the need for the exclusive use of Multi-Extension FITS (MEF) files, and fully machine readable contain of the FITS headers has lead the Elixir team to upgrade most of its interfaces (note: this also applied to NEO). This resulted overall in an increase of the data quality and their archival value.

2.3.2 Tuning the Elixir recipes

As stated earlier, the initial Elixir recipes for MegaCam were simply derived from the lesson learned with CFH12K. However, several steps were changed before the first data were made available to the PIs. Even more care was taken on deciding the recipe to adopt for the first release of CFHTLS Elixir data to CADC in January 2004 (which would end up at Terapix for stacking).

The most important improvement was the photometric flatness of the data. High quality photometric grids were obtained on high density star fields and allowed a much better sampling of the illumination function across the mosaic. The first Elixir release of January 2004 benefited from this upgrade, and further works on this after a light baffle was installed on the telescope early 2004 led to a re-release of all the Deep data since the beginning of the survey in the fall of 2004 (the highest photometric precision is needed to accomplish the SNLS scientific goals). The re-release of the Wide and Very-Wide data using the upgraded master flat-fields hasn't taken place yet.

The fringe pattern behavior on the MegaCam appears to be a function that depends not only of the sky level but also on the airmass. While we first expected the fringe residual to be in the vicinity of 0.1%, some data cannot be corrected at better than 1% under the current scheme. The implementation of a principal component analysis using master fringe frames derived from specific sky conditions is under preparation. This will lead to a re-release of all i' and z' CFHTLS data to CADC.

2.3.3 Real-Time Analysis Systems

While DADS delivers rapidly the raw data to Waimea, this is not a format adequate for the precise photometry work needed for the supernovae search. The observing configuration Telescope + MegaPrime + MegaCam turned out to be very stable from run to run and the level of photometry accuracy and cleanness of the detrending step could be achieved using the master detrending frames from the previous runs. Elixir detects automatically when a new image from the CFHTLS is available on the archive disk and immediately processes it and pushes it in a location visible to all the RTAS. Elixir processes one MegaCam image in approximately 3 minutes, which means than the SNe RTAS (this is the program that get prioritized transfer from the summit) have access to fully processed image, photometrically flat within 1% as specified by Elixir, within only 10 to 15mn after the image was acquired at the telescope.

There are currently 3 RTAS at CFHT: the French and the Canadian Supernovae clusters, and the French RGB RTAS which uses the Very Wide survey data.

2.3.4 Providing Elixir data to the CFHTLS community

In the course of 2004, it appeared that the optimal Elixir processed data (using the master detrending frames from the observing run), should be delivered as soon as possible to CADC and open to the CFHTLS community. A maximum delay of 20 days after the end of a run was set as a goal. Elixir&DADS have been able to deliver the data within this delay over the past semester, except for a couple of occurrences where glitches in the system (CFHT and/or CADC) caused a distribution to be missed. CFHT feels confident that this delay can be respected from now on.

In order to avoid the delivery and real-time processing being affected by a machine failure (Elixir cluster is composed of 4 main nodes, only one severe failure happened so far, in Nov. 2003), all machines now have a duplicate node. Swapping from a cluster configuration to another can be accomplished in a matter of minutes.

2.3.5 Pushing the limits of wide-field photometry

MegaCam uses a Sloan filter set: u*,g',r',i',z' (u* is different from the Sloan u' as the E2V detectors used in MegaCam have a better near-UV response). This forced us to use the very bright sparse photometric standards published by Sloan in 2002 (Smith et al., AJ). The scatter in the zero points noticed from observing run to observing run is larger than one would expect from the instrument and the site of Mauna Kea. The SNLS is by far the most demanding program in terms of photometric accuracy and the Canadian and French teams have put a lot of efforts in understanding the limitations of the instrument and calibration that cause a scatter of +/– 2 to 3% on the final photometry of stacked frames (result also confirmed by Terapix). Since CFHT had set a goal of a 1% photometry accuracy on the the Elixir processed data, there is currently an effort led by the SNLS teams and the CFHT Elixir team to tackle this issue. New observing strategies are now in place during each QSO run:

- Defocus (slightly) the instrument when observing primary standards
- Observe the deep fields with short exposures to build tertiary standards
- Acquisition of the photometric grid in median seeing conditions
- Review in minute details of all the steps that could affect photometry

3 Realities of observing: why the CFHTLS data rate is slower than expected

This section aims at describing the factors responsible for the slower than expected data rate of the survey compared to the initial plan which was based on statistics from previous years and performance numbers derived from the instrument design documents. As stated in the previous section, some of the following factors are also responsible for the difference in time balancing between the three survey compared to the initial goals (i.e. weather and image quality).

3.1 Weather

Bad weather on Mauna Kea is by far the first cause for lack of efficiency: either the dome is closed because the outside conditions are very bad (snow, high humidity, high wind), or the sky is fully overcast with cirrus too thick (more than 1 mag. attenuation) to allow any useful science to be conducted, or the seeing is so bad that no highly ranked scientific program can be executed (seeing higher than 1.2" in the r' band).

Here are the statistics on the time lost to weather since the official start of the CFHTLS which is approximately taken as the beginning of the semester 2003B (actual official start was on May 30th 2003). These numbers come from the QSO semester reports available on the QSO web site at CFHT (these numbers apply only for nights when MegaPrime was on telescope).

Semester	2003B	2004A	2004B
Total number of QSO nights	104	118	116
Nights lost to weather (cumulated)	31	32	25
Fraction	30%	27%	22%

However the winter nights are longer by 2.5 hours than summer nights (10h45mn versus 8h15mn) and the bad weather tends to dominate in the winter of course, hence the loss in time is even more severe overall as the survey was defined with a yearly night length average of 9h30mn. From these statistics, let us consider the total number of night lost to weather a year to 60, and than 66% of that bad weather happens during the period October to March which has an average night length of 10h30mn, and the rest during the April-September period which have an average length of 8h50mn. The total numbers of hours lost is actually $60\times0.66\times10.5 + 60\times0.34\times8.8$, which gives a total of approximately 600 hours. Using now the average number of hours per night throughout the year (9.5 hours), this is equivalent to 63 nights of observing lost to bad weather over a year. The total number of QSO nights was on average these past three semesters 113, hence a more accurate number to quantify the fraction of time lost to weather over a year is 28% $((113\times2)/63)$.

The bad seeing is also an important factor: of all the observations obtained with MegaCam since first light from the g' to the z' band, 25% exhibit an image quality higher than 0.9 arcsec., the highest acceptable image quality for the CFHTLS. While the 0.9 to 1.2 arcsec. image quality domain can still be useful for some QSO programs, this means that the fraction of clear time left after the bad weather has been accounted for is not fully usable for the CFHTLS: only about 80% or so of it since the Deep SNLS can make use of some of the degraded seeing periods, but only at the very beginning and very end of an observing run.

Overall, these recent weather statistics show that Mauna Kea suffers from unusual bad weather: 28%, far more than the 20% at most that was expected at the time the CFHTLS was being defined. And the fourth semester, 2005A started in February, is even worse with 75% of cumulated loss due to very bad weather after two QSO observing runs.

3.2 Technical problems

MegaPrime/MegaCam is a very complex instrument, and the first semester of operation was rich in failures typical for a young instrument. However, the number of hours lost to technical problems haven't decreased as quickly as they should have if they had follow the typical law of an instrument's life. Some parts of the instrument (particularly the autofocus and the wide-field corrector) have required a significant amount of sky engineering time which accounts just the same way as failures forbidding science observations.

Note that all failures were thoroughly investigated and addressed in a way that should forbid them from reoccurring. The most severe failure was a vacuum problem in the cryostat which caused the first two weeks of the first 2004A QSO run to be totally lost (and the rest was lost to the weather).

The following table, based on the QSO semester reports, gives the total number of nights lost to engineering plus technical problems:

Semester	2003B	2004A	2004B
Total number of QSO nights	104	118	116
Nights lost to E&T (cumulated)	12	24	9
Fraction	12%	20%	7%

On average, the fraction of time lost to engineering and technical problems is 13%. However, MegaPrime has proven to be very stable over the past 6 months, and now that the image quality issue of the wide-field corrector has been mostly addressed (see below), a more realistic number of 5% should be expected for the coming semesters.

Hence so far, accumulating the bad weather conditions, the engineering, and the technical problems, the fraction of time lost is 41% (versus the "planned" 25%).

3.3 Observing efficiency

Out of the remaining time available for observing, one has to consider the instrument and telescope overheads. Here is a list of the most significant overheads during a night, and the typical total contribution to the time budget per night.

Action	**Time per action**	**Budget per night**	**Goal/Status**
Exposure overhead	50 sec	70×50 = 58 mn	70×45 = 52 mn
Guide star acquisition	30 sec	20×30 = 10 mn	$<$ 2 mn
Filter change	90 sec	10×90 = 15 mn	Optimized (exclude standards)
Focus	400 sec	8×400 = 53 mn	0 sec with auto
Photometric standards	540 sec	2×540 = 18 mn	0 sec with tertiary standards
Dome rotation $> 45^o$	n/a	10 mn	Reduced with less standards

Current developments at CFHT are focused on the most significant overhead: the focusing of the instrument. The autofocus has been a very hard point with an erratic behavior but now that the image quality has been greatly improved on MegaPrime, the guiding/focus probes also benefit from the very significant improvement in image quality at the edge of the field of view. CFHT expects the autofocus to be easier to troubleshoot and implement at this point. The guide star acquisition is being made faster with a change of regime of the guide probe motors: the guide star acquisition should be 10 to 15 times faster as of April 2005. On-going work in Elixir for reducing the error in the absolute photometric calibration could lead to eventually drop the Sloan bright standards in favor of tertiary standards located in the four CFHTLS deep fields (since they are the most visited places by CFHT throughout the year).

Currently, the total number of hours coming from the instrument and telescope overheads is approximately 2.7 hours (derived from the previous table with the photometric standards included since they are not considered scientific QSO observations). This represents 28% of the average length of a night. With the improvements listed above this number should go down to 14% (1.3 hours).

It is important to recognize the fact that the Queued Service Observations mode is not meant to primarily optimize the observing efficiency of the telescope, it is meant primarily to optimize the scientific return of the telescope by ensuring that the top ranked programs get executed (for semester 2004B, the completion of A ranked programs reached 97%, and 67% for B ranked programs), and that the time balance between funding agencies is respected (this is achieved within 1 to 2% at the end of the semester). QSO will never be as efficient as an observing program which would integrate with long exposures on a couple of fields a night in a single filter. QSO has to deal every night with a large number of exposures of various lengths on many locations across the sky, and with many filter changes. The following tables show this clearly (the integration time here is open shutter time):

Semester	03B	04A	04B
Total number of QSO nights	118	116	113
Total number of nights with light integration (open night)	93	105	112
Total number of exposures	6323	5289	7879
Average number of exposures per open night	68	50	70
Maximum open shutter time over a single night (hrs)	7.60	6.76	8.27
Average open shutter time over the number of open nights (hrs)	4.2	3.0	4.9
Total number of filter changes	1392	1524	1795
Average number of filter changes per open night	14	14	16

It is interesting to take a look at the nights which achieved the highest number of hours integrated over a single night for these three semesters, as well as an average (typical) QSO night for the 2004B semester (Nov. 17 2004):

Date	Oct.21 03	Apr.26 04	Nov.13 04	Nov.17 04
Total number of hours integrated	7.60	6.76	8.27	4.39
Number of exposures	103	66	57	113
Mean etime (sec)	265	368	520	139
Number of sky regions visited	9	7	9	14
Number of filter changes	32	22	13	32
Number of standards observation (sets)	3	2	1	2

Not surprisingly, the typical QSO night of 2004B had a lot of short exposures, many filter changes, many different regions of the sky observed, whereas the best nights of the three first semesters are simply the opposite.

The time accounting done by QSO has always included 40 seconds per exposure, a time charged to the observing programs: PIs are asked to include this exposure overhead in their time request to the TACs (note that it is underestimated compared to the real value of 50 seconds, this translate as an apparent lower QSO efficiency, say a few percent). The open shutter time plus the 40 seconds overhead per exposure is referred hereafter as the "QSO integration" time. In the following table, the difference between "Total hours of QSO integration" and "Total hours of QSO science integration" is coming from the photometric standards and the snapshot programs (really bad observing conditions used anyway to gather light).

Also, the canonic validation rate expected for MegaPrime was 90% (see section 1.1 for a description of the validation process and why it can't naturally never be equal to 100%), however the past three semesters have had many periods of unstable seeing and the validation rate has been affected by up to 10% ("Queue validation efficiency" in the next table).

Adding the bad weather, the time lost to engineering & technical problems, and the instrument overheads (including the photometric standards), the total number left for light integration (open shutter time) over the past three semesters is: 9.5×0.61 - 2.7 = 3.1 hours, which for a typical night with about 70 exposures taken and the 84% validation efficiency over the past three semesters, gives approximately 3.3 hours of QSO validation time per night, 50% off from the 6.5 hours per night upon which the survey strategy had been built.

Semester	2003B	2004A	2004B	Average
Total number of observing nights	104	118	116	113
Number of nights lost (weather, technical)	43 (41%)	56 (47%)	34 (30%)	44 (39%)
Total hours of open shutter time	394	314	553	420
Total number of exposures	6323	5289	7879	6497
Total hours of QSO integration	464	373	640	492
Total hours of QSO science integration	425	335	568	443
Total hours of QSO validated time	340	282	500	374
Queue validation efficiency	80%	84%	88%	84%
Average QSO validation (hours per night)	3.2	2.4	4.3	3.3

One must be optimistic for the future as the observing overheads and the time lost to engineering and technical problems will decrease, a trend that can be seen already in the 2004B statistics. If all of the items listed in the overhead table are addressed within the semester 2005A (human resources have been affected to the project), the semester 2005B could offer an extra 1.5 hours per night. With a planned technical problems & engineering time decreased to 5%, and still considering the current rate of bad weather (28%), the open shutter time per night would then be 9.5×0.66 - 1.3 = 5.0 hours, which brings a QSO validation time of approximately 5.2 hours per night (considering 70 exposures per night, and a 90% validation rate).

If the weather returns to the canonic value of 20% derived from the 1990s' Mauna Kea weather statistics and the instrument behaves perfectly with technical problems representing less than 2% (the maximum level set by CFHT's new standards for instruments operation), the maximum level the QSO validation could reach (still considering 70 exposures per night, and a 90% validation rate) is approxi-

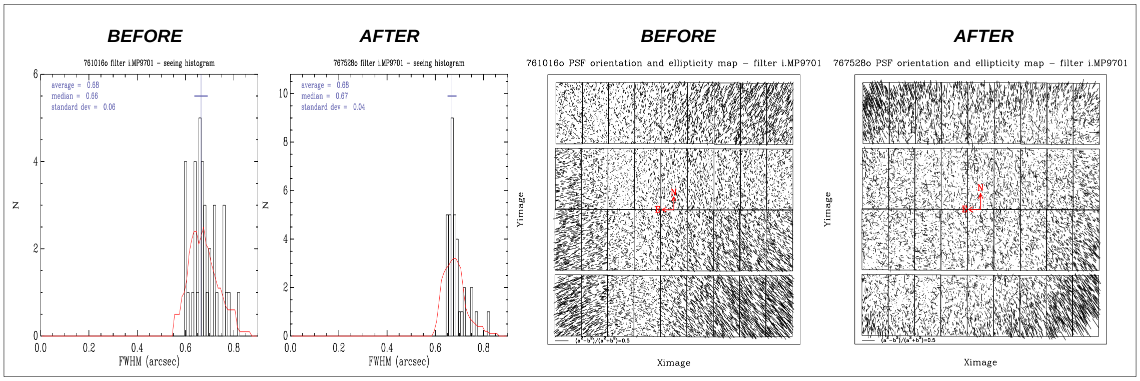

Figure 5: Illustration of the dramatic image quality improvement of the image quality before and after the flipping of lens L3. Top plot is the histogram of the image quality over the field of view, bottom plot is the star ellipticity mapping across the field of view. Plots produced by Terapix.

mately 6.2 hours per night.

In conclusion, while the semester 2005A will still run at an expected pace of QSO validation of approximately 4.5 to 5.0 hours per night, an increase up to 5.2 hours per night is accessible for semester 2005B even when adopting the poor weather statistics of the past two years. The initial goal of 6.5 hours per night is impossible to reach if the current weather conditions on Mauna Kea persist. This is a very important point further discussed in the CFHTLS mid-term review main document.

3.4 Image quality

Scientific use of MegaPrime/MegaCam started shortly after the official first light in January 2003, however the image quality proved to be far from what the specifications had called for. Over the following months, CFHT's staff converged on an optimal configuration which was judged adequate by the CFHTLS Steering Group to start integrating light on CFHTLS fields. There was a wish from the Steering Group for the configuration to be kept stable for long periods of time, say a semester or two, to allow proper calibration of the instrument characteristics and potential impact on the science goals. The CFHTLS officially started on May 30th 2003 just as the whole instrument optical configuration was frozen.

Investigations on the poor image quality delivered by MegaPrime's wide-field corrector progressed at a moderate pace until the CFHT Users' Meeting held in Campbell River (Canada) in May 2004, where it was made clear by the community (especially the Steering Group) that the optical performances of the instrument were impeding seriously the CFHTLS, even putting it at risk. A task force was put together at CFHT, leading to the analysis of all the various optical part of the instrument and the telescope (primary mirror), a very complex and risky task since all the optical elements from the wide-field corrector had to be dismantled again and again and put back together as MegaPrime's on-sky schedule could not be perturbed. Nothing obvious showed up throughout all these studies, though the biggest and largest lens of the wide-field corrector slowly appeared as the most serious suspect (L1).

It was during one of those dismantling/remounting operation (usually taking place within a single day, putting a very high pressure on CFHT's engineers and technicians) that the third lens of the wide-field corrector (L3) was mounted back flipped upside-down. L3 is indeed a very flat lens with little power and with mechanical mounts allowing the lens to be flipped, the mistake is understandable. The "error" was noticed as the image quality on the MegaCam images spectacularly improved! Figure 5 shows the difference between the normal optical configuration (based on the design, L3 is indeed not supposed to be flipped, what the flipping did is correct some aberrations caused by the lens L1, a proper modelling of this effect in undergoing at CFHT). First, one can see that the histogram of the image quality is much narrower: the image quality across the whole mosaic is a lot more uniform. Second, the shape of the PSF is a lot better with the ellipticity also almost uniform across the whole field of view. Only the top left and lower right corners now have image quality beyond the specifications, this leaves 94% of the field of view within specifications for the science drivers identified for MegaPrime/MegaCam. This configuration, in place since December 2004, has proven stable over time and the same improvement is seen in all bands, from u* to z'. The following table gives the image quality (arcsec) at the center of each of the 36 CCDs (organized in 4 rows of 9 chips) for the best image ever obtained at CFHT with MegaPrime, shortly after L3 was flipped:

0.54	0.49	0.48	0.49	0.47	0.46	0.46	0.49	0.54
0.49	0.50	0.49	0.47	0.46	0.46	0.45	0.44	0.45
0.51	0.53	0.54	0.50	0.47	0.47	0.46	0.46	0.49
0.53	0.54	0.54	0.53	0.50	0.48	0.48	0.54	0.65

Investigations on the wide-field corrector are continuing at CFHT, but at a slower pace since the new optical configuration puts the instrument very close from the design specifications and allow all of the planned science to be done, especially the CFHTLS weak-lensing component.

3.5 Image quality and time constraints

Due to the time constraints on the Deep SNLS and the Very Wide components, the QSO team has often no choice but to gather a fraction of the data under seeing conditions worse than what was requested. The inverse is also true and data are sometimes gathered for the CFHTLS is much better seeing conditions than requested. In the end, these two trends balance themselves and the average image quality over the whole data set sits within the initial specifications.

4 CFHT services tailored to inform the CFHTLS community

4.1 Global CFHTLS communication channels

The Steering Group (SG) is composed of nine members (see Figure 6) and bears the responsibility of conducting the CFHTLS for the Canadian and French communities. It is in consequence the central node of all communications. The SG reports to the SAC on a semester basis through written documents. The SG has recently decided to also send reports to the TACs since they tend to program large programs in direct competition with the CFHTLS which makes QSO scheduling very complex (CFHT on its own has emitted such wishes to the TACs, always unsuccessfully though). As stated earlier, a constant communication channel is open between the SG (the coordinators) and the QSO team throughout the year. The Data Oversight Group (DOG) is in charge in making sure all the data are properly formatted and communicate with the relevant entities when a problem arises.

To this day, the CFHTLS community is composed of more than 200 scientists, all with a clear attachment to Canadian and French institutions. Communication to this community by the SG is primarily through email (see below), and regular workshops and meetings.

Since it is important to show to the rest of the world the advances on the CFHTLS, both CFHT and Terapix have a very open policy for accessing www pages full of information on the status (progress, data quality, etc...) of the survey. This is today, along with many oral presentations given throughout the year (50 talks related to the CFHTLS have been given in 2004, mostly by SG members), the principal

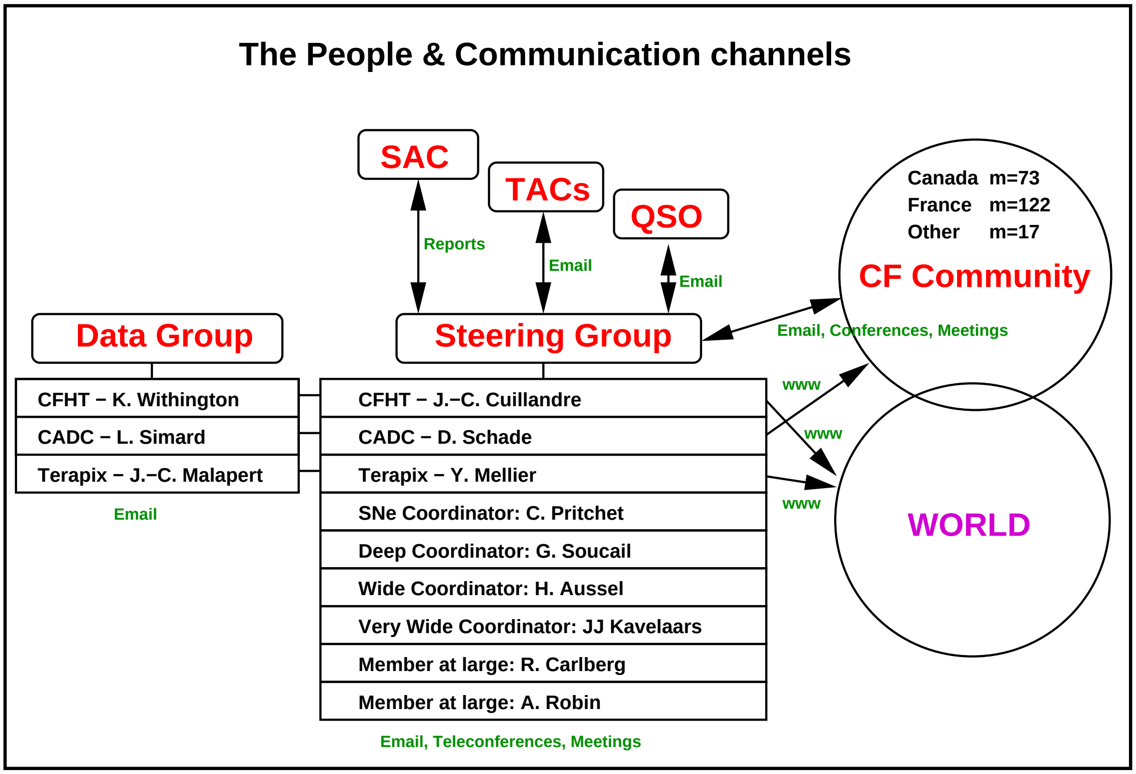

Figure 6: Communication channels for the CFHTLS

mean used to create an interest for the CFHTLS as the time of the worldwide 2006 release approaches.

4.2 CFHTLS mailing lists

The CFHTLS mailing lists are hosted and managed by CFHT. When a scientist is interested in joining the CFHTLS, a submission form is submitted to the CFHT Executive Director who, based on the current home institution, grants, or not, the right to mail to, and receive from, the various CFHTLS mailing lists. This also gives access to the CFHTLS mail archive at CFHT, and of course access as well to the CFHTLS Elixir and Terapix data product at CADC (the CADC provides a different login and password).

4.3 Main CFHTLS web site

When the CFHTLS officially started, a new web site focused on the status and progress of the survey replaced the previous web site which was focused on the definition of the survey. All the important information has of course been ported to the new CFHTLS web site, which is recognized at the “official” CFHTLS web site. The web site is maintained by CFHT’s Steering Group member and is regularly updated with the latest news. The goal is to keep track of all the events related to the CFHTLS.

“www.cfht.hawaii.edu/Science/CFHTLS/”

4.4 CFHTLS Data web site

A different web site keeps track of all the observations obtained by CFHT for the CFHTLS. It is basically a database open to everyone, constantly kept up to date. Status of the observation can be browsed either as a whole since the beginning of the survey (and provides for example the global statistics which are used by the CFHTLS coordinators to check on the balance between the three surveys), or on a hourly basis during an observing run (this is especially useful for supporting the Real Time Analysis Systems). To keep every piece of information presented on this web site clear, all parameters and quantities are described in details.

"www.cfht.hawaii.edu/Science/CFHTLS-DATA/"

4.5 MegaPrime/MegaCam web site

Still based on the same web design to keep a familiar look&feel, the MegaPrime/MegaCam site presents all the technical information relevant to the use of the data by a scientist. Its primary function is indeed to educate on the process involved in operating the instrument, to get acquainted with its properties to submit a time proposal (exposure time calculator, etc...), along with a special focus on the data pre-processing and calibration by Elixir. Due to the crucial importance of keeping track of the instrument history (failures, change in properties, etc...) throughout its life in regards of its impact on the data, all events that are relevant to the data (failure of an amplifier, change in the optical configuration) are recorded on the web site. This is of great importance for the CFHTLS which timescale approaches the lifespan of the instrument.

"www.cfht.hawaii.edu/Instruments/Imaging/MegaPrime/"

4.6 Scientific meetings

Starting 2004, the Steering Group members have also organized yearly national meetings for each individual C&F community. Also, this year CFHT is organizing in collaboration with the IAP the first CFHTLS scientific meeting. It will be held in May 2005 in Paris at the IAP. The goal is to gather together the Canadian and French communities to offer them an opportunity to share their experience with the CFHTLS data and their first scientific analysis.

"www.cfht.hawaii.edu/LSW05/"

5 Conclusion

The CFHTLS operation under the NOP at CFHT is a success: the instrument MegaPrime/MegaCam is more and reliable; QSO produces quality data in abundance and following the requested specifications, especially the time constraints; Elixir supports the Real Time Analysis systems, produces rapidly quality data that can be streamed easily into the CADC archive where they made available to the CFHTLS community, and then straight into the Terapix stacking pipeline; DADS delivers promptly all the CFHT data products to CADC: the Elixir processed images (FITS) and the associated ancillary elements. The CFHTLS community can follow all the advancement of the survey, as well as the status of the instrument, on a daily or monthly basis by consulting web sites hosted and maintained at CFHT. Overall the CFHTLS has set new standards in the NOP chain which are a direct benefit to the standard PI programs.

However, MegaPrime/MegaCam is not yet sufficiently efficient on the sky with overheads which are still too high and do not allow the gathering of science data at the rate initially expected. CFHT is currently putting some effort into this now that the previous main issue has been addressed: MegaPrime now delivers excellent image quality across the full MegaCam field of view.

More worrisome, the weather conditions on Mauna Kae over the past two years have been worse than usual and have caused great damages on the observing statistics. It is unlikely at this point that the initial predicted QSO validation rate of 6.5 hours per night can ever be achieved, but 5.5, or even 6.0, hours per night is within reach if the overheads are eventually reduced. This lack of efficiency (mostly due to the weather) has had a serious impact on the advances of the CFHTLS which was designed with a QSO validation rate of 6.5 hours per night whereas on average the first three semesters have brought only 3.3 hours per night.