

IceCube Instrumentation and Online Systems

The IceCube Collaboration

1 Introduction

1.1 IceCube Science

(1 page)

1.2 A Functional Description of the IceCube Instrument

In order to observe astrophysical neutrinos, the primary science goal of the experiment, IceCube exploits the fact that charged particles moving through the ice at super-luminal speed emit Cherenkov photons. An enormous detection volume is required since the cross-sections of neutrinos are small for producing secondary charged particles in interactions with ordinary matter. The glacial ice cap at the South Pole is about 3km thick and therefore predestined as operation site since it is not only offering aplenty interaction material but also a medium with unmatched high quality. Cherenkov light is produced in cascades of neutrino-induced muons penetrating the deep, other high-energy particles as well as generated by by cosmic-ray muons that result from cosmic-ray interactions in the atmosphere above Antarctica. Due to a Cherenkov photon yield of $\mathcal{O}(10^5)$ visible photons per GeV of shower energy, the long optical attenuation length in South Pole ice and large-area photomultipliers (PMTs) it is possible to instrument cubic kilometers of ice with a rather wide spacing of detectors. The basic detection unit in IceCube in order to capture the Cherenkov light is the digital optical module or *DOM* which is covered in great detail in Sec. ???. Encapsulated in a 1/2" thick glass pressure sphere to withstand the extreme pressure in the deep ice, the main components of a DOM are a 10" PMT, embedded high-voltage generation, a flasher calibration board, and a mainboard containing the analog and digital processing circuitry for PMT pulses. The digitized data is fed to a central computing facility at the surface via a unique cable system, see Sec. ??. Aspects of detector deployment and ice drilling are covered in Sec. ??. An overview of the data flow as well as its readout, processing and filtering are subjects of Sec. ?? where we also cover the data handling, monitoring and operational performance of the observatory. The IceCube instrument consists of three sub-detectors – IceCube, DeepCore and IceTop – using the same instrumentation design of embedded digital optical modules and associated surface readout. A schematic layout of the array is shown in Fig. 1

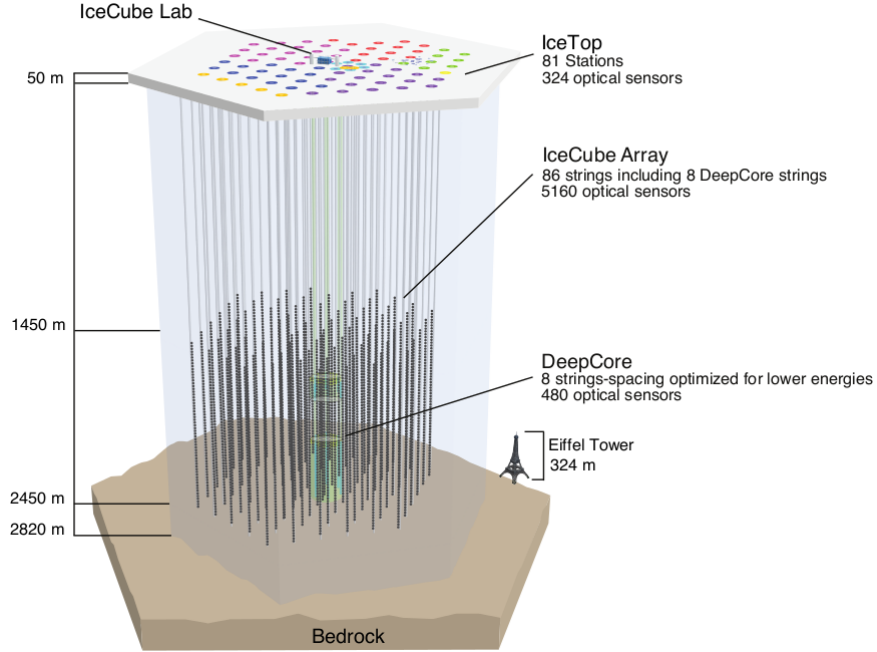


Figure 1: The IceCube Neutrino Observatory with its sub-array DeepCore and the air shower array IceTop.

32 1.2.1 IceCube

33 In order to detect the Cherenkov photons emitted by charged particles traversing the ice,
 34 5160 DOMs are deployed between 1450 m and 2450 m below the glacial surface on 86
 35 vertical strings each holding 60 DOMs deployed along a copper cable. The main *deep-*
 36 *ice* array consists of 78 strings with a vertical separation of the DOMs on each string of
 37 17 m. These strings are deployed on a hexagonal grid with 125 m horizontal spacing and
 38 spanning a volume of one cubic kilometer of ice. *Talk about energy range covered by this*
 39 *instrumentation and primary science goal*

40 1.2.2 DeepCore

41 The remaining subset of in-ice DOMs is deployed in the deep ice below a depth of 1750 m
 42 forming a denser instrumented volume. This sub-array, the DeepCore [1] consists of eight
 43 specialized and closely spaced strings of sensors located around the central IceCube string.
 44 Its inter-string spacings of 75 m and inter-DOM spacings of 7 m are optimized for the
 45 detection of atmospheric neutrinos with energies typically in the range from 100 GeV to
 46 400 TeV [3].

47 **1.2.3 IceTop**

48 The air shower array IceTop [4] consists of 81 Cherenkov tanks filled with clear ice that
49 are arranged in pairs on the same, approximately 125 m, triangular grid on which the
50 in-ice array is deployed. The two tanks at each surface station are separated from each
51 other by 10 m. Each tank contains two standard IceCube DOMs. Air showers initiated in
52 the atmosphere by cosmic rays are typically spread over a number of stations. The light
53 generated in the tanks by the shower particles (electrons, photons, muons and hadrons) is
54 a measure of the energy deposit of these particles in the tanks.

2 The Digital Optical Module

(Chris Wendt; 10 pages)

2.1 A Functional Description of the DOM

The DOM (Figure ??) is the fundamental data acquisition unit for IceCube, containing a downward-facing 10" diameter photomultiplier tube (PMT) and associated circuit boards that allow near-autonomous operation. Data acquisition, control, calibration, communication and low voltage power conversion are integrated in one annular circuit board that fits around the neck of the PMT (Main Board, Section ??) [?]. Separate circuit boards generate PMT high voltage, interface to the PMT pins [?], delay PMT signals, and generate calibration light flashes that can reach other DOMs. Key requirements for the DOM include the precise recording of a wide variety of PMT pulse widths and amplitudes, robustness in a challenging deployment environment, and long term reliability.

The PMT detects signals from neutrino events ranging over energies 10 GeV–10 PeV and distances from a few meters to 500 m away. Corresponding PMT waveforms can have amplitudes from 1 mV up to and beyond the linearity limit of the PMT (~ 2 volts) and widths from 12 nsec to 1500 nsec. In order to accommodate such a variety of signals, the DOM includes multiple digitizers with overlapping dynamic range and different sampling speeds [?]. Each DOM is triggered independently by detection of individual photons, starting a recording of the PMT waveform that includes further photons arriving up to 6.4 μ sec later. The trigger time is saved along with the waveform shape, which reveals the times of arriving photons relative to this reference. The DOM typically accumulates such triggered data for a period of ? to ? sec before sending as a block.

DOMs transmit their data to computers in the IceCube Laboratory building using a twisted wire pair that also provides power. Wire pairs are bundled together to form the vertical down-hole cables and the horizontal surface cables. Each wire pair is shared between two DOMs, with data transfers initiated by a surface computer. Separately, dedicated wiring to neighbor DOMs above and below allows quick recognition of local coincidences where nearest or next-to-nearest neighbors trigger within a common 1 μ sec time window. Local coincidence triggers often have complex PMT waveforms reflecting multiple photons detected in each DOM, which are therefore saved in full detail; otherwise the DOM saves abbreviated information appropriate to single photon detection [?].

The DOM is capable of interpreting commands from the surface that specify tasks for configuration, data taking and transmission, monitoring or self-calibration. Self-calibration functions establish PMT and amplifier gains as well as sampling speed. The RAPCal system [?] is implemented for tracking each local DOM clock's offset from universal time, allowing PMT pulses that were independently recorded in many DOMs to be recombined into events by surface computers.

92 2.2 Components

93 2.2.1 Glass Sphere and Harness

94 The glass sphere housing has diameter 13" and thickness 0.5". Spheres are specified to
95 protect the inside electronics and PMT against long term applied pressure of 250 bar (2.6 km
96 water depth) as well as temporary overpressure up to 690 bar during refreezing of melted
97 ice in the drill hole. They were produced by Benthos, Inc., based on a design for deep sea
98 environments but using glass with very low potassium or other radioactive trace elements
99 that would contribute to the dark noise count rate (Section ??). Optical transmission was
100 measured for representative samples as 93% at 400 nm, decreasing to 50% at 340 nm and
101 10% at 315 nm (normal incidence, excluding Fresnel reflection).

102 Each sphere is assembled from two hemispheres that mate precisely at the equator; after
103 evacuating and backfilling with dry nitrogen, a butyl rubber sealant is applied around the
104 seam, and covered with wide plastic tape. The interior gas pressure is set to 0.5 bar so the
105 seal remains tight even at ambient south pole air pressure 0.6 bar.

106 The DOM is held by an aluminum band with rubber gaskets against the glass above
107 and below the equator seam. Figure ?? shows how it is attached to the main down-hole
108 cable via a system of steel rope and chain that carries the weight load around the DOM.
109 The main cable bends around the DOM, and the DOM axis stays vertically aligned with
110 the string.

111 2.2.2 Cable Penetrator, Cable and Connector

112 A customized penetrator assembly brings three wire pairs out through a 16.3 mm hole in
113 the DOM glass. They are routed inside the umbilical cable, visible in Figure ??, and
114 terminate at a pressure-tight, waterproof connector that mates with a similar connector
115 and thus continues each pair into the main cable. One wire pair connects ultimately with a
116 computer in the IceCube Laboratory building, carrying power and the bidirectional digital
117 communications stream. The other two pairs lead to neighbor DOMs directly above and
118 below (Figure ??), carrying local coincidence digital pulses that signify time correlated hits
119 in nearby DOMs (Section ??).

120 The penetrator itself is a customized stainless steel feedthrough, with a plastic shell that
121 is also molded onto the umbilical cable jacket. The penetrator has an o-ring seal facing the
122 glass outside and is secured by a nut inside the sphere. External mechanical features like
123 the penetrator are subject to large stresses during deployment and the refreezing process;
124 the right angle bend was chosen for robustness, based on previous experience deploying
125 AMANDA modules. The complete assembly, including umbilical cable and type XSJJ
126 connector, was produced by SEACON (California).

2.2.3 PMT, Gel and Magnetic Shield

DOMs use the 10" diameter Hamamatsu R7081-02 PMT, or the corresponding high-quantum-efficiency (HQE) version for Deep Core strings. Its properties have been measured and described in [?]. It is specified by Hamamatsu for the wavelength range 300 nm–650 nm, with peak quantum efficiency around 25% (34% for HQE) near 390 nm. It features a box-and-line dynode chain with 10 stages, and is operated at gain 10^7 (Section 2.4.1).

The PMT bulb faces downwards in the bottom glass hemisphere, secured in high-strength silicone gel to a depth surrounding the photocathode area. The gel provides mechanical support for the whole assembly of PMT and circuit boards, as well as good optical coupling. Gel thickness between PMT envelope and glass sphere is approximately 1 cm. Originally the gel was supplied as General Electric RTV6136-D1, and later as a similar formulation from Quantum Silicones (Virginia, USA). It is optically clear with transmission 97% at 400 nm, 91% at 340 nm, and 65% at 300 nm (normal incidence). The refractive index is 1.41, yielding less than 0.1% Fresnel reflection as light passes from the sphere glass into the gel and then into the PMT envelope. The characteristics of the cured gel are specified to remain stable in the temperature range -70°C to 45°C .

To reduce effects of the ambient magnetic field (550 mG, 17° from vertical), a mu-metal cage surrounds the PMT bulb up to the neck join. It was constructed as a wire mesh with typical wire spacing 66 mm and wire diameter 1 mm, blocking about 4% of the incident light. The resulting interior magnetic field is a factor 2.8 below the external field, pointing mostly along the axis and therefore reducing efficiency by less than 2% for this type of PMT [?].

Other interior DOM components are held in place by attachment to the PMT, mostly via screws into a molded plastic collar glued around the neck. The PMT base board is soldered directly to the pins.

2.2.4 Main Board and Delay Board

The Main Board and its operation has been described in detail in [?]. It interfaces to other boards as shown in Figure ?? and itself provides many key functions of the DOM:

- Control all the devices inside the DOM, including the high voltage power supply for the PMT, the flasher board, and various sensors (pressure, temperature, power supply voltage monitor). Also supply necessary DC power to the subsystems.
- Digitize the PMT waveforms, using the custom ASIC (ATWD: analog transient waveform digitizer) and a continuous sampling ADC.
- Carry out computing functions. This includes executing PMT gain calibration, compressing the digitized waveform, temporarily storing the data, creating data packets and time stamping them.

- 163 • Communicate with the data acquisition system on the surface.
- 164 • Exchange timing pulses with the surface DAQ to calibrate the internal DOM clock.
- 165 • Exchange "local coincidence" pulses with the adjacent DOMs.

166 PMT waveforms are captured by the ATWD chips with sampling period 3.3 nsec, start-
 167 ing after a discriminator trigger is recognized by the sequencing logic. The total sampling
 168 interval is 427 nsec, encompassing the time spread of most photons arriving from up to
 169 100 m away. The Delay Board is connected between the discriminator and the ATWD
 170 preamplifiers, adding a pretrigger window in this recorded waveform that allows precise
 171 fitting of the leading edge time. It contains a ~ 10 m long, 0.25 mm wide, serpentine cop-
 172 per trace embedded in the dielectric and sandwiched between ground planes, giving a total
 173 delay of about 75 nsec. After accounting for triggering delays, this causes the waveform
 174 recording to start at least 10 nsec before the leading edge time.

175 2.2.5 HV Supply and Divider

176 The PMT high voltage subsystem consists of a resistive voltage divider circuit (PMT
 177 Base) directly solder-mounted on the PMT and a separate high voltage control board. The
 178 high voltage control board includes a DAC and an ADC for setting and reading out the
 179 PMT high voltage, connected to the Main Board with a digital interface. It also holds
 180 the high voltage generator, which is a custom encapsulated module designed by EMCO
 181 (California). The maximum high voltage is 2047 volts, specified for up to $30 \mu A$ current.
 182 The set voltage is proportional to the DAC output, and the actual voltage is monitored via
 183 a high-impedance divider and the ADC. The output ripple is less than 1 mV, and stability
 184 is better than 200 ppm over 8 hours. Power consumption is 300 mW at full load.

185 The generator output is carried to the PMT Base Board via a high voltage coaxial
 186 cable. This board is described in [?]. Its voltage divider presents a total resistive load
 187 of $130 M\Omega$. The PMT is operated with cathode at ground potential, so the anode signal
 188 output is AC coupled using a 1:1 bifilar wound toroid transformer mounted on the Base
 189 Board. The transformer secondary is then wired to the Main Board analog input with a
 190 coaxial cable.

191 2.2.6 Flasher Board

192 Each IceCube DOM contains a flasher board. The standard IceCube flasher board, which
 193 is included in every DOM except the color DOMs described below, is a circular board fitted
 194 with 12 LEDs (ETG-5UV405-30) with a peak wavelength of 399 nm; the FWHM of the
 195 LED spectrum is 14 nm. The LEDs are arranged in pairs, evenly spaced around the board
 196 with a 60° separation between each pair. One LED in each pair points outward horizontally,
 197 the other LED is tilted upward at an angle of 48° , which is close to the Cherenkov angle

in ice ($n = 1.36$). The angular emission profile of each LED has a FWHM of 30° in air, which is modeled as a Gaussian emission profile with $\sigma = 13^\circ$. After refraction through the DOM glass and into the ice, the value of σ is 9.7° in the polar direction and 9.8° in the azimuthal direction for the tilted LEDs, and 9.2° in the polar direction and 10.1° in the azimuthal direction for the horizontal LEDs.

The LEDs are controlled by a current pulse applied to each LED through a high speed MOSFET driver with a series resistor. The LEDs can be turned on individually or in any combination of the 12, using a bitmask in the flasher DAQ configuration. The photon output of the LED depends on the width and brightness, or amplitude of the driving current pulse. The current amplitude is controlled by the brightness setting in the flasher DAQ, which may have a value between 0 and 127; the maximum LED current is 240 mA. The current pulse width is controlled by the width setting in the flasher DAQ, which is twice the current pulse width in nanoseconds and may have a value between 0 and 127; the maximum current pulse width is 70 ns. The minimum stable pulse width that we can achieve is about 6 ns. The photon output is related to the brightness and width by the following empirically derived relationship:

$$N = 1.17 \times 10^{10} (0.0006753 + 0.00005593B) \left(W + 13.9 - \frac{57.5}{1 + B/34.4} \right) \quad (1)$$

where N is the number of photons, B is the brightness setting (maximum value 127) and W is the width setting (maximum value 127). The photon output per LED ranges from 10^6 to 10^{10} photons per flash, which is equivalent to the energy output of GeV - 100 TeV cascades. The maximum LED pulse rate is 610 Hz. The current waveform is read out by ATWD digitizer channel 3 during flasher operation in order to measure the precise rise time of the flasher pulse. The flasher light output begins 8.3 ns after the current pulse is recorded. Although flashers can be operated in multiple DOMs in the same run, the DAQ does not support time-synced flashing of LEDs on different DOMs, so coincident flasher events happen only by chance.

The flasher LEDs are used for a variety of calibration purposes:

- verifying the timing response of the DOMs throughout the software
- measuring the position of the deployed DOMs in ice
- measuring the optical properties of the ice
- verifying the performance of cascade reconstruction algorithms in measuring position, direction and energy

Color DOMs

There are 16 DOMs (8 on string 79 in the center of IceCube, 8 on string 14 on the edge of IceCube) fitted with multiwavelength flasher boards, called color DOMs or cDOMs.

Table 1: Properties of the cDOM LEDs

LED	nominal λ	measured λ	σ air	σ DOM, polar	σ DOM, azimuthal
UVTOP335-FW-TO39	340 nm	338 nm	51.0°	36.1°	42.9°
NS370L_5RFS	370 nm	371 nm	55.2°	39.1°	42.9°
LED450-01	450 nm	447 nm	6.8°	4.8°	5.3°
B5-433-B505	505 nm	494 nm	6.4°	4.5°	4.9°

Each cDOM includes 3 LEDs with a nominal wavelength of 505 nm, 3 LEDs with a nominal wavelength of 450 nm, 3 LEDs with a nominal wavelength of 370 nm and 3 LEDs with a nominal wavelength of 340 nm. The LEDs are arranged in pairs as on the standard flasher board, but all LEDs point outward horizontally. The arrangement of the pairs is shown in Figure XXcdom sketch from cdom wikiXX. The properties of the LEDs are given in Table 1.

2.3 Production and Testing

2.4 Calibration

2.4.1 DOMCal

2.4.2 Flasher Calibrations

2.5 Performance and Reliability

We have over N DOM years in ice. What can be said about the reliability? This section could be quite important.

2.5.1 Baseline Stability

The digitizer pedestals are set to about 10% of the maximum value in order to capture signals that go below the baseline. The baseline refers to the average value of this offset, whereas the pedestal pattern refers to the bin-to-bin variation of the pedestal setting. The pedestal pattern is subtracted off inside the DOM, but the baseline is subtracted during data processing at the surface. Since 2012, the average baseline value is set by the DAQ configuration in order to ensure stability. The average value differs for each digitizer channel in each DOM, ranging from XX to XX ADC counts. The pedestal pattern is computed at the beginning of each run by comparing 25 averaged pedestals. The autocorrelation coefficient between the pairs of averaged pedestals is computed to detect light contamination in the pedestals, and the shift between the baseline of the pairs is calculated to determine that the baseline is stable. This procedure ensures that fewer than 1 DOM in 1000 runs will contain a contaminated baseline.

258 The baselines for each digitizer channel in each DOM are monitored with beacon hits,
 259 forced triggers which are collected at a rate of XX Hz in each DOM. The average value of
 260 the beacon baselines are very stable, with an average stability of XX counts year to year.
 261 The baselines are sensitive to radio frequency interference (RFI). In 2009, RF signals from
 262 a radar transmitter broadcasting at 46.3 MHz appeared as sinusoidal or spiky signatures
 263 in the waveform baselines. Also in 2009, a DOM 68-42 “Krabba” was damaged by a too-
 264 high voltage setting during DOM calibration, and appeared to begin sparking, causing
 265 sinusoidal waveforms to appear in the baselines of neighboring DOMs.

266 **2.5.2 Optical Sensitivity Stability**

267 The detector response in IceCube is calibrated with low energy muons as described in
 268 [?]. The detector response is monitored in each run using the track detection probability
 269 (TDP) calculated from high multiplicity muon tracks with more than 30 hits in IceCube.
 270 The muon tracks are reconstructed using the likelihood methods described in [?]; charge
 271 and time information from the DOM under study excluded from the reconstruction. The
 272 TDP is defined for each DOM as the ratio of the number of detected tracks within 100 m
 273 of the DOM to the total number of tracks within 100 m of the DOM. This ratio depends
 274 both on the optical properties of the ice near the DOM and the optical efficiency of the
 275 DOM. We do not attempt to separate these effects in the TDP, but rather use the TDP
 276 to monitor the overall stability of the detector response. Figure XX shows the TDP on
 277 string 80, which includes both standard and HQE DOMs; the TDP is 20% - 25% higher for
 278 HQE DOMs than for neighboring standard DOMs. The TDP is stable to within 1% since
 279 2012, when the baselines were stabilized by being set in the DAQ configuration. Figure
 280 XX shows the difference in the TDP for all DOMs between a run in 2012 and a run in
 281 2015.

282 The detector response stability is also measured with the *in situ* light sources in IceCube.
 283 Both the in-ice calibration laser [?] and the flasher LEDs show less than 1% difference in
 284 the total charge collected between 2012 and 2015.

285	3	The Cable Systems
286	3.1	Surface Cables
287	3.2	Surface-to-DOM Cables
288	3.3	Surface Junction Boxes

5 Online Systems

(John K; 12-15 pages)

The IceCube online systems comprise both the software and hardware at the detector site responsible for data acquisition, event selection, monitoring, and data storage and movement. As one of the goals of IceCube operations is to maximize the fraction of time the detector is sensitive to neutrino interactions (“uptime”), the online systems are modular so that failures in one particular component do not necessarily prevent the continuation of basic data acquisition. Additionally, all systems are monitored with a combination of custom-designed and industry-standard tools so that detector operators can be alerted in case of abnormal conditions.

5.1 Data Flow Overview

The online data flow consists of a number of steps of data reduction and selection in the progression from photon detection in the glacial ice to candidate neutrino event selection, along with associated secondary data streams and monitoring. An overview of the data flow is shown in Fig. ??.

Since the majority of photons detected by the DOMs are dark noise, a first-level *local coincidence* (LC) is formed between neighboring DOMs deployed along the same cable, using dedicated wire pairs within the in-ice cable. DOM-level triggers, or *hits*, with corresponding neighbor hits are flagged with the LC condition, while hits without the condition are compressed more aggressively. The LC time window as well as the span of neighbor DOMs up and down the cable can both be configured, with standard settings of a $\pm 1\mu\text{s}$ coincidence window and neighbor span of 2 DOMs.

All DOM hits are read out to dedicated computers on the surface by the data acquisition system (DAQ). The next level of data selection is the formation of *triggers* by the DAQ system. LC-flagged hits across the detector are examined for temporal and in some cases spatial patterns that suggest a common causal relationship. A number of different trigger algorithms run in parallel, described in Sect. ?. All hits (both LC-flagged and non-LC hits) within a window around the trigger are combined into *events*, the fundamental output of the DAQ, and written to disk. The event rate is approximately 2.5 kHz but varies with the seasonal atmospheric muon flux, and the total DAQ data rate is approximately 1TB/day.

The DAQ also produces *secondary streams* that include time calibration, monitoring, and DOM scaler data. The scaler data, which is monitoring the noise rate of each DOM in 1.6 ms bins, is used in the supernova data acquisition system [?] to detect a global rise from many $O(10)$ MeV neutrino interactions occurring in the ice from a Galactic core-collapse supernova. The time calibration and monitoring streams are used to monitor the health and quality of the data-taking runs.

The raw DAQ event data is then processed further with a number of *filters* in order to select a subset of events (less than 10%) to transfer over satellite to the Northern Hemi-

sphere (see Sect. ??). Each filter, typically designed to select events useful for a particular physics analysis, is run over all events using a computing cluster in the IceCube Lab. Because of limitations both on total computing power and bounds on the processing time of each event, only fast directional and energy reconstructions are used. The processing and filtering system is also responsible for applying up-to-date calibrations to the DAQ data; processed events, even those not selected by the online filters, are stored locally for archival.

A dedicated system for data movement handles the local archival storage to tape or disk, as well as the handoff of satellite data (see Sect. ??). This includes not only primary data streams but also monitoring data, calibration runs, and other data streams.

[Add experiment control paragraph]

Figure: data flow diagram indicating major subsystems to be described in this section: DAQ (incl. SNDAQ), PnF, I3Live, and SPADE/JADE.

5.2 SPS and SPTS

South Pole System: breakdown of computing hardware used at the pole between hubs, DAQ, PnF, other machines, and infrastructure. Internal network bandwidth. Redundancy, system monitoring (Nagios), and paging system.

Brief mention of SPTS as northern test and validation system. Replay capabilities.

5.3 Data Readout and Timing

5.3.1 Communications and Cable Bandwidth

Description of communications protocol and messaging strategy. Reference to RAPCal and how it fits in. Event compression.

5.3.2 Master Clock System

The GPS clock and time string fanout tree, from master clock (and hot spare), to Tier I and Tier II fanouts, the DSB card, and into the DOR card.

5.3.3 DOR Card and Driver

DOR card description: comms / readout, power control and measurement, RAPCal initiation, and clock string readout. Clock modes (internal / external). DOMs per card and cards per hub.

Brief description of driver. Proc file interface. Data transfer over PCI bus via DMA.

Figure: Combined clock fanout tree hierarchy and hub diagram (DSB and DOR cards, power distribution).

5.4 Processing at the Surface

(Dave G; 2-3 pages)

IceCube’s data acquisition system (DAQ) is a set of components running on dedicated servers in the IceCube Lab. As described in ??, physics data is read from the DOMs by the StringHub component and a minimal representation of each HLC hit is forwarded to the “local” Trigger components. The “local” Trigger components run these minimal hits through a configurable set of algorithms and form windows around interesting temporal and/or spatial patterns. These time windows are collected by the “global” Trigger and used to form non-overlapping trigger requests. These trigger requests are used by the Event Builder component as templates to gather the complete hit data from each StringHub and assemble the final events.

5.4.1 DOMHub and Hit Spooling

The *StringHub* is responsible for periodically reading all available data from each of its connected DOMs and passing that data onto the downstream consumers. It also saves all hits to a local “hit spool” disk cache, as well as queuing them in an in-memory cache to service future requests from the *Event Builder* for full waveform data.

The StringHub component is divided into two logical pieces. The front-end *Omicron* controls all of the connected DOMs, forwarding any non-physics data to its downstream consumers and sorting the hits using an HKN1 tree before passing them to the back-end *Sender*. The Sender caches SLC and HLC hits in memory, then forwards a few fields from each HLC hit to the local Triggers.

After the Trigger components have determined interesting time intervals, the Event Builder will send each interval to the Sender, and it will return a list of all hits within the interval, pruning the in-memory hit cache of all older hits after each interval.

In order to avoid saturating the 2006-era network, the StringHub sends only simplified HLC hit data to the local Triggers. Fields extracted from the full HLC are the string and DOM, the time of the hit, and a few DOM trigger fields (type, mode, configuration ID), for a total of 38 bytes per hit.

Dave G doesn’t know the details of this

One core assumption of the DAQ is that each component operates on a time-ordered stream of data. IceCube’s DAQ uses its *Splicer* to accomplish this. The Splicer is a common object which gathers all input streams during the setup phase; no inputs can be added once it’s started. Each stream pushes new data onto a “tail” and the Splicer uses the HKN1 algorithm to collate the data from all streams into a single output stream. When a stream is closed, it pushes an end-of-stream marker onto its tail which causes the Splicer to ignore all further data.

The HKN1 (Hansen-Krasberg-1) algorithm at the core of the Splicer uses a tree of *Nodes* to sort a constant number of time-ordered data streams. Each Node contains a

398 queue of pending data, a handle for a peer Node, and another handle for a sink Node
399 (shared with the peer) to which sorted data is sent.

400 The tree is assembled by associating each input stream with a separate Node and adding
401 the Node to a list. This list is then run through a loop which pulls off the first two Nodes,
402 links them together as peers and creates a new Node which will act as the sink for both
403 Nodes. The new Node is then added to the list. The loop exits when there is only one
404 Node left.

405 When a new piece of data arrives, it is added to the associated Node's queue. That
406 Node then goes into a loop where it compares the top item from its queue with the top
407 item from the peer's queue and the older of the two items is passed onto the sink Node
408 (which adds the item to its queue and starts its own comparison loop.) This loop stops as
409 soon as the original Node or its peer has no more data.

410 Along with physics data, DOMs produce three additional streams of data. The scaler
411 data, which is monitoring the noise rate of each DOM in 1.6 ms bins, is used in the su-
412 pernova data acquisition system [?] to detect a global rise from many $O(10)$ MeV neutrino
413 interactions occurring in the ice from a Galactic core-collapse supernova. The time calibra-
414 tion and monitoring streams are used to monitor the health and quality of the data-taking
415 runs.

416 As hits move from the front end to the back end, they are written to the "hit spool", a
417 disk-based cache of files containing hits. These files are written in a circular order so that
418 the newest hits overwrite the oldest data, and the first hit time for each file is stored in an
419 SQLite database.

420 Subsystems such as SnDAQ and HESE can send time interval requests to a HitSpool
421 master daemon which then gathers the hits in that interval from each hub and forwards
422 them to a final "sender" daemon which bundles them up and submits them to JADE for
423 transfer to the North where the full waveforms can be used for further analysis.

424 One problem with the current DAQ design is that it only reads data when the DAQ
425 is running, so the detector is essentially "off" during hardware failures or the periodic full
426 restarts of the system. There is a plan to split the StringHub into several independent
427 pieces to eliminate these blind spots. The front end (Omicron) piece will be moved to an
428 always-on daemon which continuously writes data (including secondary, non-physics data
429 and other metadata) to the disk cache. Part of the back end (Sender) piece will become a
430 simple HitSpool client which reads data from the disk cache and sends it to the downstream
431 consumers, while another simple component will listen for time intervals from the Event
432 Builder and send back lists of hits taken from the hit spool.

433 5.4.2 Supernova System

434 SN secondary stream from DOMs and SNDAQ reference. Interface to hitpooling.

435 5.4.3 Triggers

436 General description of trigger architecture. Separation of trigger window and readout
437 window. How trigger windows depend on geometry. Thorough escription of all different
438 trigger algorithms. Trigger and readout window merging.

439 Table: standard settings for triggers

440 Figure: trigger windows and readout windows.

441 Figure: example bright multi-trigger event.

442 Figure: SLOP triplet geometry?

443 5.4.4 Event Building

444 Readout requests to StringHub components and packaging of waveforms into events. Spool-
445 ing to disk and interface to PnF.

446 5.4.5 Configuration

447 Tree of XML configuration files for components, triggers, and DOM settings.

448 5.4.6 Distributed Network Control

449 CnC server description. XML-RPC.

450 5.5 Online Filtering

451 (*Erik B.; 3-4 pages*)

452 5.5.1 Overview

453 The online processing and filtering system is charged with the immediate handling of all
454 triggered events collected by the data acquisition system. This treatment includes appli-
455 cation of calibration constants, application of event characterization and filtering software,
456 extracting data quality monitoring information, generation of realtime alerts for events of
457 astrophysical interest and creation of data files and metadata information for long term
458 cataloging. The online processing and filtering system is a custom software suite that uti-
459 lizes a computer cluster of ~ 20 standard servers located in SPS computing cluster at the
460 experimental site at South Pole, Antarctica. The online processing and filtering system has
461 been in operation since the start of operation of the 22 string configuration of IceCube in
462 2007.

463 In IceCube, each triggered event consists of a collection of digitized waveforms recorded
464 by the digital optical modules (ref daq-dom paper). To be useful for physics, each of these
465 waveforms requires application of calibration constants that allow the waveform units to
466 be converted from the raw units (ADC counts per sample bin) to more physical units (mV

measured in each fixed time, \sim ns bin). These calibration constants are independently measured (domcal ref?) and stored in an online database for use by the online processing and filtering system. Next, each DOM's waveform is deconvolved using the known DOM response to photons to extract the light arrival time and amplitude information. This series of time and amplitude light arrival information for each DOM is the base until for event reconstruction and characterization. The online processing and filtering system encodes this information for each DOM in a compact data format known as the SuperDST, and occupies \sim 5% of the file size of the full waveform information. Any DOM readout whose SuperDST information is found not to a good representation of the original waveform, or sees very high amounts of light also have the full waveform readout saved in addition to the SuperDST.

Each event is then characterized with a series of event reconstruction algorithms that attempt to match the observed patterns of recorded light in the SuperDST with known patterns of light from track and showering event hypotheses (ref reco papers? e-reco paper?). These characterizations (location, direction, and energy) and their overall goodness-of-fit produced by these reconstructions are used to select interesting events by a filter selection. The filter criteria are set by the IceCube collaboration for each season and are tuned to select events of interest to specific analyses. Each season there are about 2 dozen unique filter selections in operation. Some of these filters trigger are designed to search for events of wide astrophysical interest to the scientific community and trigger alerts that are distributed to followup observatories worldwide.

The online processing and filtering system also extracts and aggregates data quality and monitoring information from the data as is it processed. This information includes stability and quality information from the DOM waveform and calibration process, rates of DOM readouts, and rates and stability information for all detector triggers and filters. This information is aggregated for each data segment and reported to the IceCube Live monitoring system.

Finally the online processing and filtering system writes several data files that make up the long-term data catalog of the IceCube experiment. These include:

- *Filtered data files* These files contain only events selected by the online filter selections. These events generally only include the SuperDST version of the DOM information and results from the online event reconstructions. These files are queued for transmission to the IceCube data warehouse by the data handling system using the TDRS satellites.
- *SuperDST data files* These files contain the SuperDST version of DOM readout information for all triggered events as well as summary information from the online filtering process. This file set is intended as the long-term archive version of IceCube data.
- *Raw data files* These files contain all uncalibrated waveforms from all DOMs for every

506 event. This large data set is saved until final data quality assurance on the SuperDST
507 sample can be completed.

508 (Q: include some information on file catalog sizes per day?)

509 5.5.2 System Design

510 The online processing and filtering system uses a modular design, where each module is
511 responsible for a portion of data processing built around a central master server node
512 and a scalable number of processing client nodes. The central master server focuses on
513 data distribution and aggregation tasks (requesting data blocks from the DAQ, collating
514 event monitoring information and writing data files), while the client process focus on the
515 per-event processing tasks (event calibration, reconstruction, analysis, and filtering).

516 The system is built upon the IceCube analysis software framework, IceTray (ref?),
517 allowing standard IceCube algorithms to be used in the online processing and filtering
518 system without modifications. Additionally, the system uses the Common Object Request
519 Broker Architecture (CORBA) system as means for interconnecting the modular portions
520 of the system. Specialized classes are used to provide CORBA interconnections within the
521 IceTray system, allowing file-like interfaces that let data to stream from one component
522 to another using native IceTray formats. Use of a CORBA name server and a dynamic
523 architecture allow for the addition and remove of filtering clients as needed to meet the
524 processing load from annual filtering changes and overall rate variations from seasonal
525 variations in the detector trigger rate.

526 5.5.3 Components

527 The flow of triggered event data in the online processing and filtering system is shown in
528 Figure 2 highlighting the flow of data from the DAQ system, through the master server
529 and clients, to files on disk and online alerts. Several standard components in the online
530 processing and filtering system include:

- 531 • *I3DAQDispatch* is a process to pickup event data from the data acquisition system
532 data cache and forward to the PFServer components.
- 533 • *PFServers* are central data flow managers within the online processing and filtering
534 system. These servers receive data from the *I3DAQDispatch* event source, distribute
535 events to and record results returning from the *PFClient* farm, and send filtered
536 events to file writer, online monitoring and alert components. Typically there are 4
537 servers used in operation.
- 538 • *PFClient* is the core calibration, reconstruction and filtering process that is applied
539 to each triggered event. In normal operation, ~ 400 of these clients operate in parallel
540 to filter events in real time.

- *GCDDispatch* is a DB caching system to prevent the 400 PFClient processes from overwhelming the DB system when requesting calibration information. This system aggregates DB requests, makes a single database query and shares the results with all PFClients.
 - *PFWriters* are responsible for creation of files and meta-data for the IceCube data catalog. These files are written in standard IceTray file format. There is one writer component for each file type created.
 - *PFOnlineWriter* is responsible for extracting alert information from the data and forwarding these alerts in real time to the IceCube Live system.
 - *PFMoniWriter* is responsible for aggregating per-event monitoring information, creating histograms and forwarding them to the IceCube Live monitoring system.
 - *PFFiltDispatch* and FollowUp clients are responsible for looking for bursts of neutrino events on timescales from 100 seconds up to 3 weeks in duration. Any significant burst of neutrinos found generates alerts sent to partner observatories worldwide.
- (Q: remove discussion of GFU, OFU, and PFFiltDispatch since this is being moved north?)

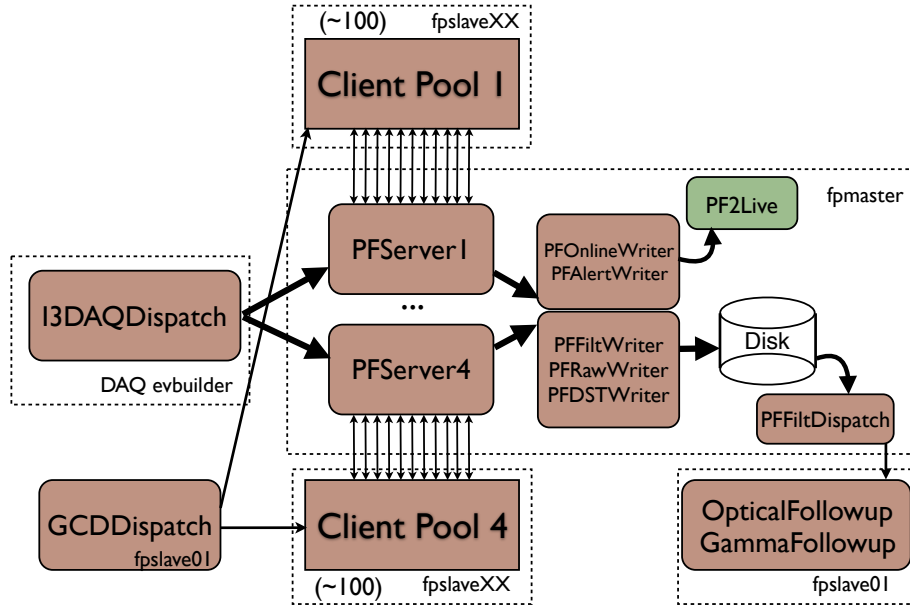


Figure 2: Internal components of the Online Processing and Filtering System. Arrows highlight the flow of data within the system.

557 5.5.4 Performance

558 The online processing and filtering system is designed to filter triggered events as quickly
 559 as possible after collection by the data acquisition system. A key metric is processing
 560 system latency, defined as the duration of time between the data acquisition trigger and
 561 the completion of event processing and filtering. A typical latency history for the system
 562 is shown in Figure 3, showing typical system latencies of ~ 20 seconds.

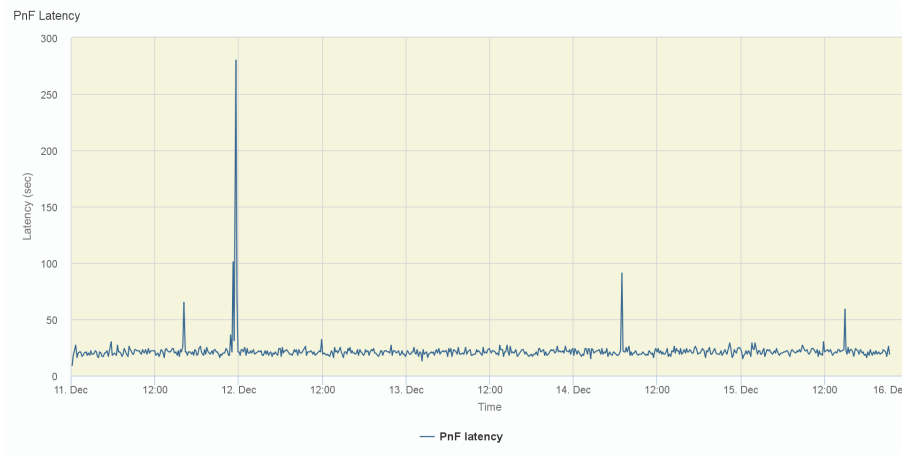


Figure 3: Typical Online Processing and Filtering System latency for a several day period. The latency defined as the time between DAQ event time and time when the online filtering processing is complete. The spikes in latency correspond to DAQ run transitions.

563 (Q: include other system performance information? Performance bottlenecks?)

564 5.6 Data Handling

565 (*P. Meade; 1 page*)

566 *Generation description of system architecture. Stream definitions, dropboxes, and data*
 567 *pickup. Archival vs. transfer to TDRSS system.*

568 Data handling is provided by three servers named jade02, jade03, and jade04. The jade
 569 servers operate independently of one another and each of them are capable of handling the
 570 nominal data volume by itself. Having three servers allows for data handling to continue
 571 seamlessly in case of hardware failure or maintenance.

572 Each server runs a copy of the Java Archival and Data Exchange (JADE) software
 573 (stylized “jade”). As its name implies, the jade software is written in the Java programming
 574 language. It is a reimplement and expansion of earlier prototype software called
 575 South Pole Archival and Data Exchange (SPADE), written by Cindy Mackenzie. The
 576 jade software has four primary tasks: consumption, archival, satellite transmission, and
 577 real-time transmission.

578 The jade software is configured with a number of data streams, which consist of a data
579 server, a dropbox directory, and a filename pattern. The data stream dropbox directories
580 are checked on a regular basis for new files. A file pairing scheme (binary and semaphore)
581 prevents files from being consumed before they are finished being produced. For each file,
582 a checksum calculated on the data server is compared to a checksum calculated on the jade
583 server. This method ensures that the file was copied without error. After this, the original
584 data file is removed from the data host.

585 After consumption, files are routed according to the configuration of their data stream.
586 Files that are too large to send via the satellite link are archived to a configurable number
587 of archival media copies. The prototype SPADE software archived to LTO tapes, while
588 the later jade software archives to large (2+ TB) hard disk drives. All of the archival data
589 is buffered on the jade server until the archival media is complete. In case of failure while
590 creating the archival media, all of the files can be immediately written to fresh archival
591 media with a single command.

592 Files that are too large to send via the real-time link, but small enough to send via the
593 satellite link are queued for satellite transmission. The jade software attempts to bundle
594 multiple files together into 1 GB bundle archives to allow satellite link operators to manage
595 the daily data transmission. Very large files (≥ 1 GB) are split apart into multiple 1 GB
596 bundles for the same reason. The jade software will only transfer a configurable number
597 of bundles to the satellite relay server. If satellite transmission is not possible, the jade
598 software will buffer the excess bundles on the jade server, to avoid flooding the relay server
599 unnecessarily.

600 Small files (≤ 50 KB) with high priority status information are sent via the real-time link.
601 The real-time link is provided by the IceCube Messaging Service (I3MS). The jade software
602 uses JeroMQ, a pure Java implementation of the ZeroMQ (ZMQ) protocol, to connect to
603 I3MS. In cases where the real-time link is not available, I3MS will queue the messages to
604 be sent when the link becomes available. All I3MS messages are also sent to jade to send
605 via the satellite link to ensure delivery if the real-time link should be unavailable for an
606 extended period of time.

607 5.7 IceCube-Live and Remote Monitoring

608 IceCube operations are controlled and monitored centrally by IceCube-Live, a suite of
609 high-level software implemented mostly in the Python language. IceCube-Live consists
610 of two major components: LiveControl, responsible for controlling data taking operations
611 and collecting monitoring data, and the IceCube-Live website, responsible for processing
612 and storing monitoring data as well as presenting this data in webpages and plots that
613 characterize the state of the IceCube detector.

614 5.7.1 LiveControl

615 LiveControl is the central control component of IceCube operations. The program executes
616 in the background as a daemon and accepts user input through XML-RPC. Operators typ-
617 ically enter commands and check the basic detector status using a command-line interface.
618 LiveControl is responsible for controlling the state of DAQ and online-processing, start-
619 ing and stopping data-taking runs, and recording the parameters of these runs. Standard
620 operation is to request a run start, supplying a configuration file specifying the DOMs
621 to include in data taking. LiveControl then records the run number, configuration, start
622 time, etc. and sends a request for DAQ to begin data taking. After data taking commences
623 successfully, LiveControl waits a specified amount of time, generally eight hours, then stops
624 the current run and automatically starts a new run using the same configuration. This
625 cycle continues until stopped by a user request or a run fails. In case of failure, Live-
626 Control attempts to restart data taking by starting a new run. Occasionally a hardware
627 failure occurs, and it is impossible to start a run with the supplied configuration because
628 requested DOMs are unpowered or temporarily unable to communicate with the IceCube
629 DAQ. In this case, LiveControl cycles through predefined partial-detector configurations
630 in an attempt to exclude problematic DOMs. This results in taking data with half of the
631 IceCube strings or fewer, but it greatly reduces the chance of a prolonged complete outage
632 where no data at all is recorded.

633 A secondary function of LiveControl is the collection, processing, and forwarding of
634 monitoring data from DAQ, online-processing, and other components. Monitoring data
635 consists of a JSON dictionary with a well-defined format including a creation time, sender
636 name, priority, data name, and either JSON data or a single integer or floating-point value.
637 This data is forwarded to LiveControl using ZeroMQ and queued internally for processing.
638 A few monitoring quantities indicate serious problems with the detector, e.g. the online-
639 processing latency is too high. LiveControl provides a database of checked monitoring
640 values indexed by service and data name and raises an alert if the value is out of the
641 specified range or hasn't been received in a specified amount of time. The alert usually
642 includes an email to parties responsible for the affected subsystem and, for serious problems,
643 triggers an automated page to winterover operators. Several other types of monitoring data
644 trigger a response by LiveControl. These include alerts generated internally by subsystems,
645 and such alerts may trigger emails and pages from LiveControl. All monitoring data are
646 forwarded to the IceCube-Live website for further processing and display.

647 5.7.2 IceCube-Live Website

648 Two operational copies of the IceCube-Live website exist: one inside the IceCube network
649 at the South Pole, fed directly by LiveControl using ZeroMQ, and one in the Northern
650 Hemisphere fed by several message transport systems depending on the priority of the par-
651 ticular monitoring data, summarized in table 2. IceCube generates approximately 300,000
652 monitoring messages daily. These compress to ~ 100 MB, are sent over JADE, and are

Priority	Transport System	Daily Messages	Daily Data	Typical Latency
1	ITS	10,000	1 MB	1 minute
2	I3MS	150,000	5 MB	1–5 minutes
3	JADE	300,000	100 MB	1 day

Table 2: Statistics for IceCube monitoring messages

typically available in the Northern Hemisphere within 24 hours. Messages of priority level 2 or higher are additionally sent over the Iridium RUDICS link using I3MS, and messages of priority 1 and higher are sent over the Iridium ITS link.

Messages reaching the website are ingested by the DBServer daemon, processed, and inserted into one of several SQL database tables or a MongoDB database depending on content. Messages also may contain directives requesting DBServer to send email, specifying recipients and content, or requesting that the monitoring message be published using ZeroMQ PUB-SUB so it can be ingested by an external process. The IceCube-Live website itself uses the Django framework and contains pages that create sophisticated views of monitoring data stored in either SQL or MongoDB. These pages include a front page that displays active alerts and plots of event rates and processing latencies from the previous few hours and a page for each run that displays start time, stop time, and other essential data. The run page contains low-level diagnostic data that includes e.g. charge histograms, digitizer baselines, and occupancy for each DOM, and is used to diagnose problems with detector components during the run and to determine if the run can be used in physics analysis.

Finally, the IceCube website transmits messages using ITS and I3MS to LiveControl running at the South Pole. This capability is used to connect the popular Slack chat service with the IceCube-Live website running at the South Pole, allowing the IceCube winterover operators to chat with experts in the Northern Hemisphere during periods with no geosynchronous satellite connectivity. Southbound messages are also used to trigger the HitSpool service by the request of experts in the Northern Hemisphere if an interesting event (e.g. a SNEWS alert) occurs.

5.8 Operational Performance

(John K.; 1 page)

Explanation of how design choices, system monitoring, and winterovers result in high uptime. Discussion of median downtime and various causes of downtime. Possible basic failure analysis of hardware components.

Figure: DAQ full uptime and clean uptime percentage.

Figure (optional): Downtime histogram.

6 Outlook

Discuss Gen2 here.

References

- [1] R. Abbasi, Y. Abdou, T. Abu-Zayyad, M. Ackermann, J. Adams, J. Aguilar, M. Ahlers, M. Allen, D. Altmann, K. Andeen, et al. The design and performance of IceCube DeepCore. *Astroparticle physics*, 35(10):615–624, 2012.
- [2] R. Abbasi, Y. Abdou, T. Abu-Zayyad, J. Adams, J. Aguilar, M. Ahlers, K. Andeen, J. Auffenberg, X. Bai, M. Baker, et al. Calibration and characterization of the IceCube photomultiplier tube. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 618(1):139–152, 2010.
- [3] R. Abbasi, Y. Abdou, T. Abu-Zayyad, J. Adams, J. Aguilar, M. Ahlers, K. Andeen, J. Auffenberg, X. Bai, M. Baker, et al. Measurement of the atmospheric neutrino energy spectrum from 100 GeV to 400 TeV with IceCube. *Physical Review D*, 83(1):012001, 2011.
- [4] R. Abbasi, Y. Abdou, M. Ackermann, J. Adams, J. Aguilar, M. Ahlers, D. Altmann, K. Andeen, J. Auffenberg, X. Bai, et al. IceTop: The surface component of IceCube. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 700:188–220, 2013.
- [5] R. Abbasi, M. Ackermann, J. Adams, M. Ahlers, J. Ahrens, K. Andeen, J. Auffenberg, X. Bai, M. Baker, S. Barwick, et al. The IceCube data acquisition system: Signal capture, digitization, and timestamping. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 601(3):294–316, 2009.