Data sharing and data policies

Erik Blaufuss - University of Maryland MANTS - September 2014 - Geneva (IceCube)



Goals of this talk



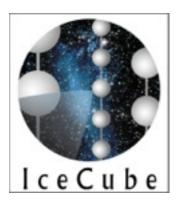
- Highlight current and planned data sharing and data release plans for different "products":
 - Raw/low level data
 - Neutrinos
 - Alerts
 - Others...
- Stimulate discussions
 - Include several open issues for discussion.
 - Will try to keep this relatively short to allow time for discussion



Product I: Raw Data



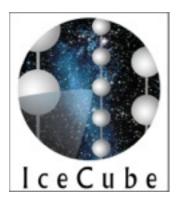
- IceCube
 - Funding agency-level requirement for release of all data
 - Publish datasets based on Level I filter for all events: trigger information and (x,y,z,t,Q) for all DOMs
 - well-defined byte-structure of simple types, with minimal support python-based toolset
 - No reconstructions, no simulation, 10's of TB/yr
- ANTARES/KM3NeT
 - Part of larger deep-sea observatories: share data (events and monitoring data) widely within consortium
 - Ex: Sea-biology research on bioluminescence
 - Wider (public) sharing is something under discussion.



Product 2: Neutrino data



- Value-added data sets, typically keyed off a specific paper/neutrino selection
 - Higher purity neutrino samples, with reconstructions, energy estimations, resolutions
 - Intended for external scientists.
- IceCube: IC22/IC40/HESE samples widely available on web
 - Tabular summary of key information for each event
 - IC59 track selection sample being used in testing with AMON system

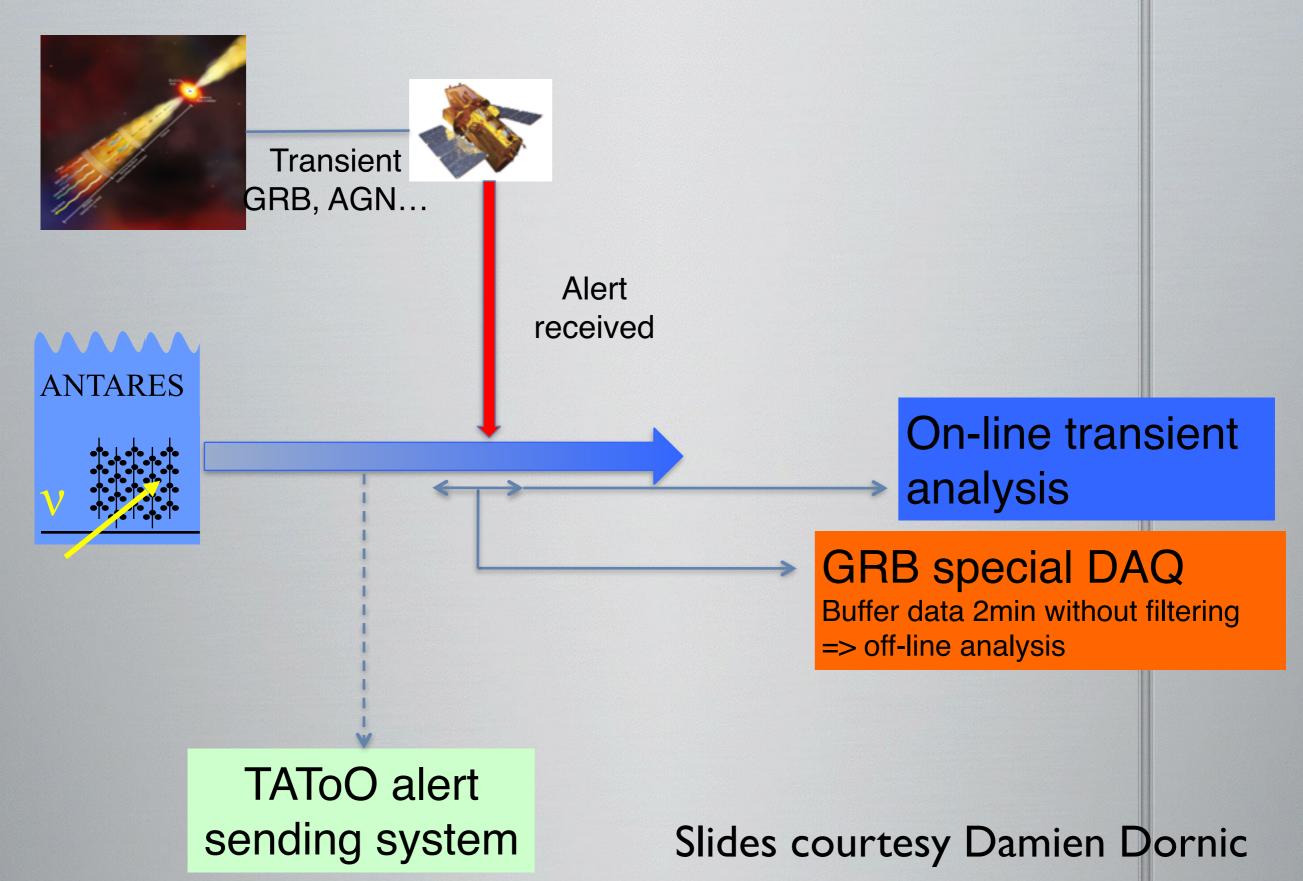


Product 3: Alerts



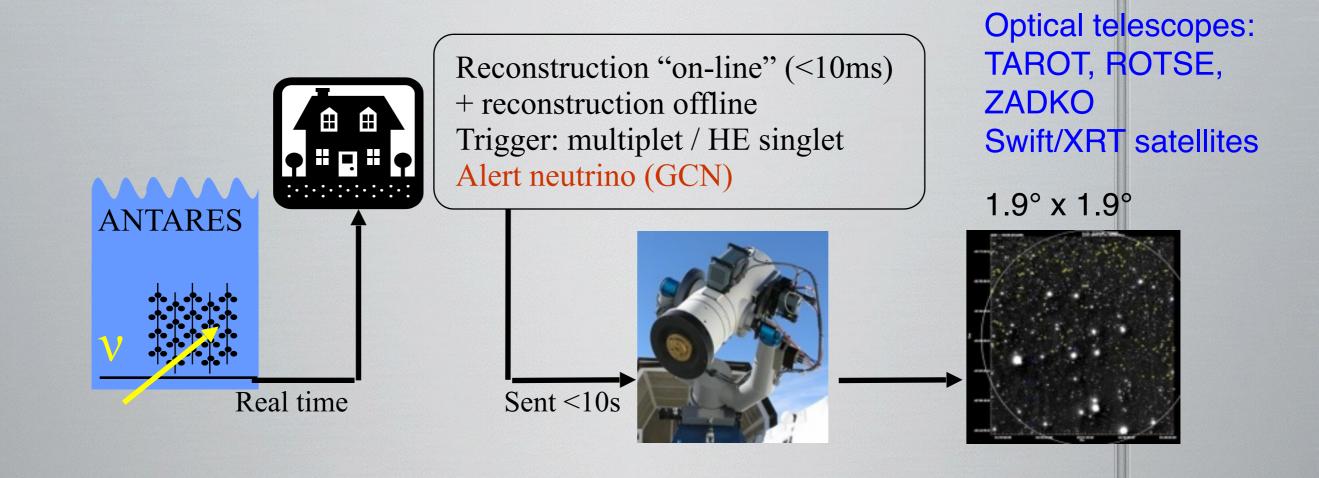
- Alerts to trigger followup observations
 - Neutrinos trigger others
 - Ex: IceCube OFU/GFU programs, ANTARES TaToO
 - Others trigger neutrino for higher sensitivity analysis
 - Ex: GCN triggers in ANTARES
- Key effort for multi-messenger searches
 - Flaring signal in other messengers in time with a (highly probable) signal neutrino would be revolutionary.

On-line framework in Antares





~6 yr experience in the on-line physics analysis



<u>3 types of triggers:</u>

- doublet (0.04/yr)
- single HE (12/yr)
- single + loc galaxy directions (12/yr)

140 alerts sent to the optical telescopessince 20095 alerts sent to Swift/XRT in 1 yr

Slides courtesy Damien Dornic

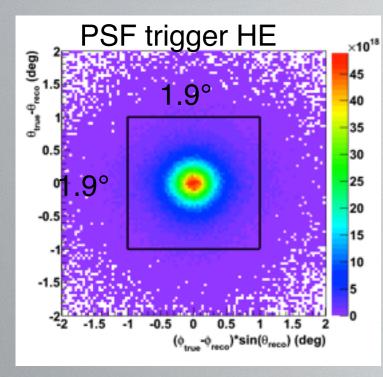
Performances alert sending

Timing performances:

- Online reconstruction + trigger: \sim 3-5 s
- Alert sending: $\sim 1-10$ s depending on the telescope response
- Telescope slewing: ~1-5 s

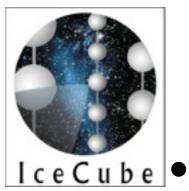
Minimum delay between the 1^{st} image and the neutrino: ~20 s

Angular performances:



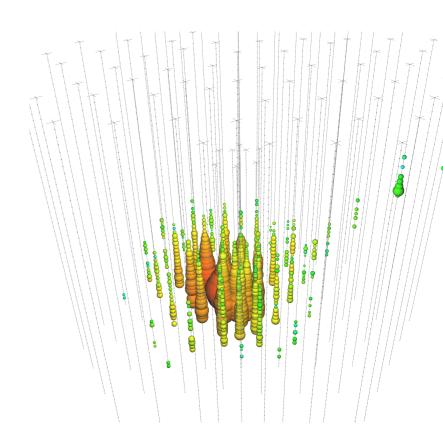
Trigger	Angular resolution	Fraction events in fov	Muon contamination	Mean energy
HE	0.25-0.3°	96% (GRB) 68% (SN)	<0.1%	~7 TeV
Directional	0.3-0.4°	90% (GRB) 50% (SN)	~2%	~1 TeV

Slides courtesy Damien Dornic

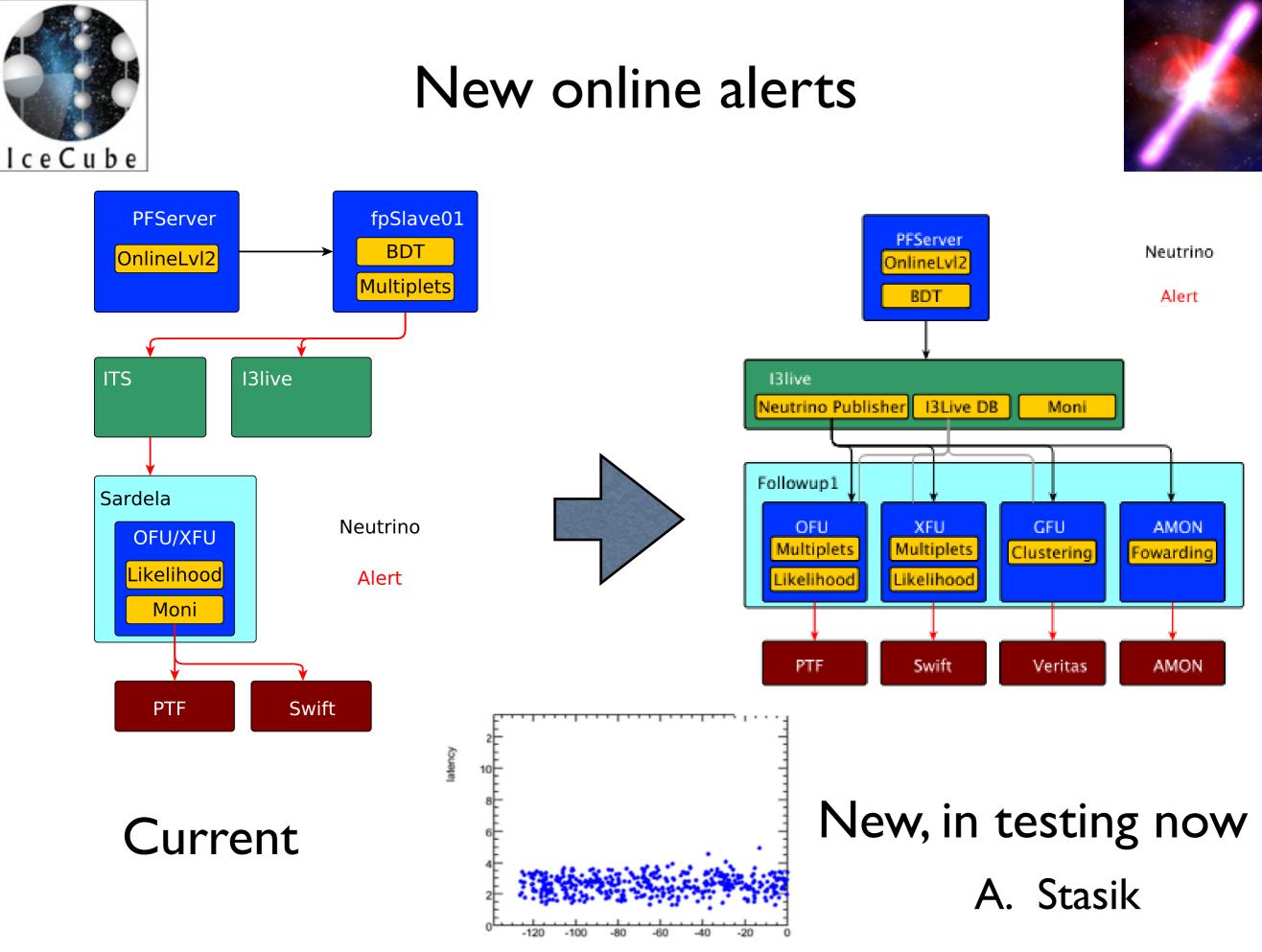


Alerts: IceCube Future

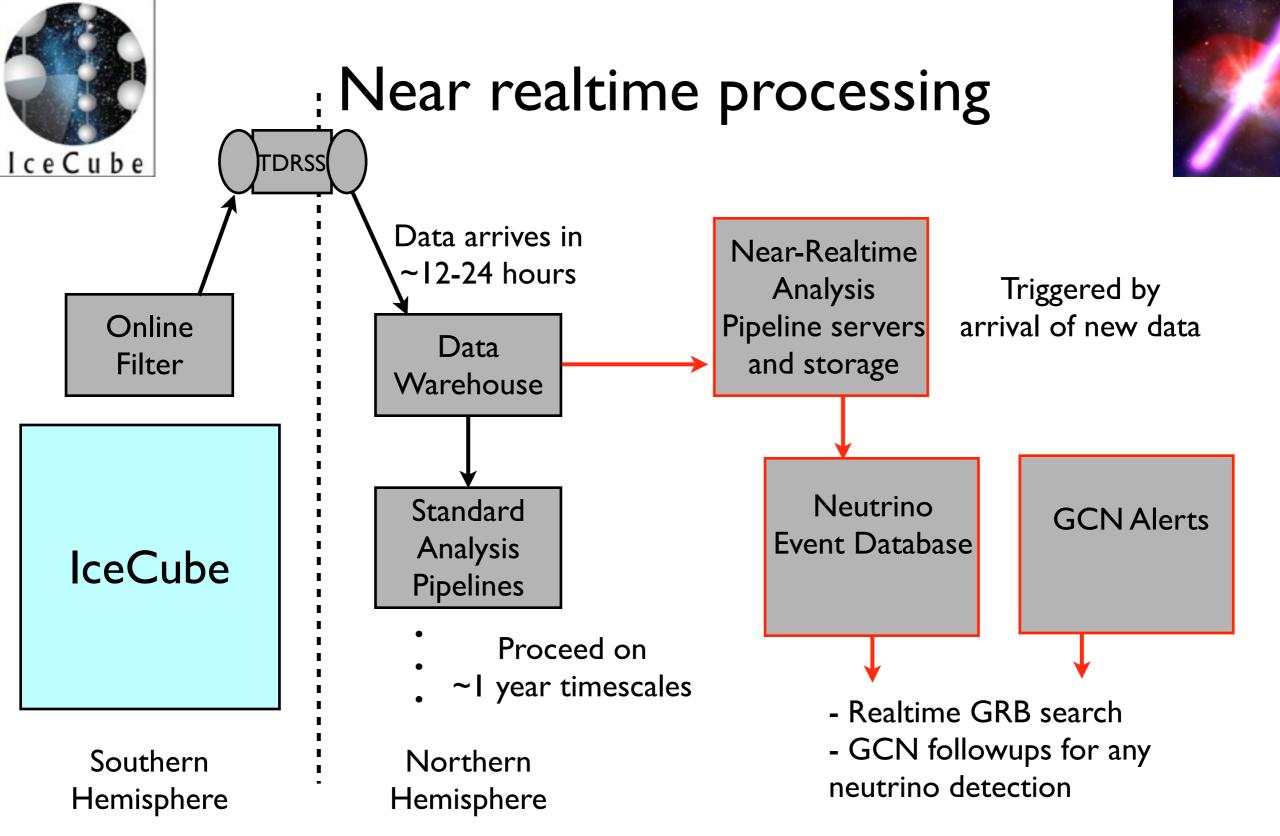
- In IceCube we are moving toward a system of public alerts for astrophysical neutrinos
 - Lots of interest in the community for these
 - Our selections are robust, and we have some reasonable estimates of how likely each event is to be signal/background
 - HESE events
 - High energy single track events
 - For each alert:
 - Date/time
 - Position on sky
 - Angular uncertainty
 - Some measure of signal/background likelihood
 - Perhaps capped at some significance



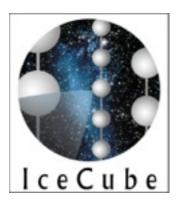




lookback time in h



- HESE alerts, GRB followup, flaring AGN searches
- Able to test systems here before moving to Pole for lower latency



IceTop 2 PeV

■ -3×10⁻³ to 3×10⁻³

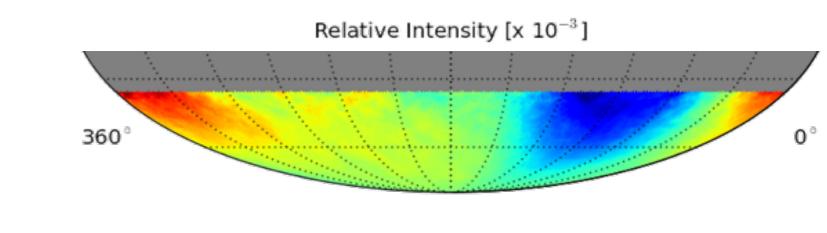
Product 4: Others



• These detectors produce interesting science beyond high-energy neutrinos

F. McNal

- IceCube: IceTop cosmic ray shower maps
 - small/large scale anisotropy in Southern hemisphere
- MOU: combination with HAWC maps of Northern Hemisphere.



F. McNally



Open discussions...



- Wider, public sharing of data or alerts is something new for the neutrino telescope community
 - In IceCube, HESE discovery has helped
 - Before an extraterrestrial signal was found, people worried about others making 1st discovery
 - Now focus has shifted to understanding what we are seeing, and multi-messenger observations could really help



Issues and concerns...



- Publication credits arising from public alerts/data.
 - Credit will be given where credit is due. They're still the "Fermi Bubbles"...
- External discoveries made in response to public alerts/ data.
 - This would be great, especially for the case of new observations triggered by an alert
- "Claims" made using public data/alerts
 - I think these should be straightforward to counter using our detailed knowledge
 - Ex: Our data is not all V_e contrary to recent claims



Issues and concerns (2)...



- When to release data?
 - Need to allow time for our own analysis. How long should this be?
- Which data formats for sharing data?
 - Supporting .i3 file format for wider audience likely more than we can support
 - Look to a more common data format
 - Ex: LIGO Open Science Center: HDF5 (also .gwf)
 - Fewer software support issues
 - Generally we have no support for a "science support center".

