



The Zwicky Transient Facility And SkyPortal/Fritz

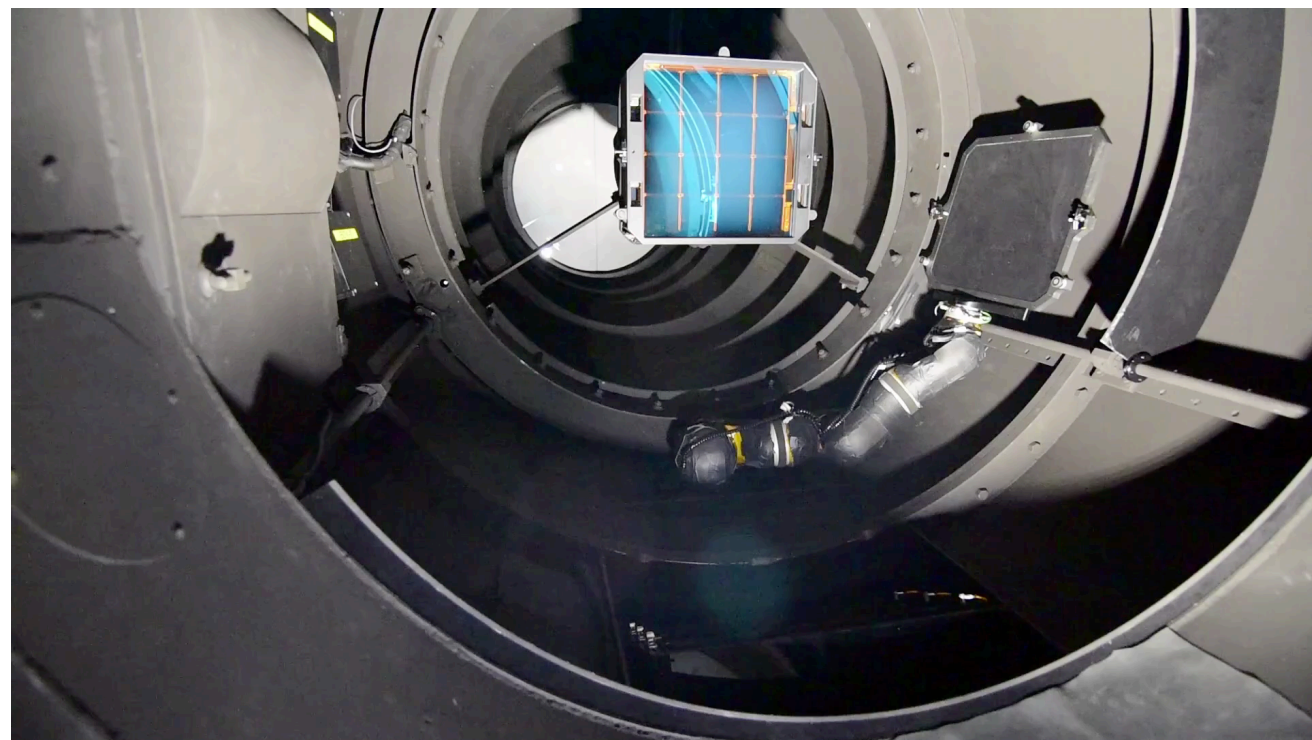
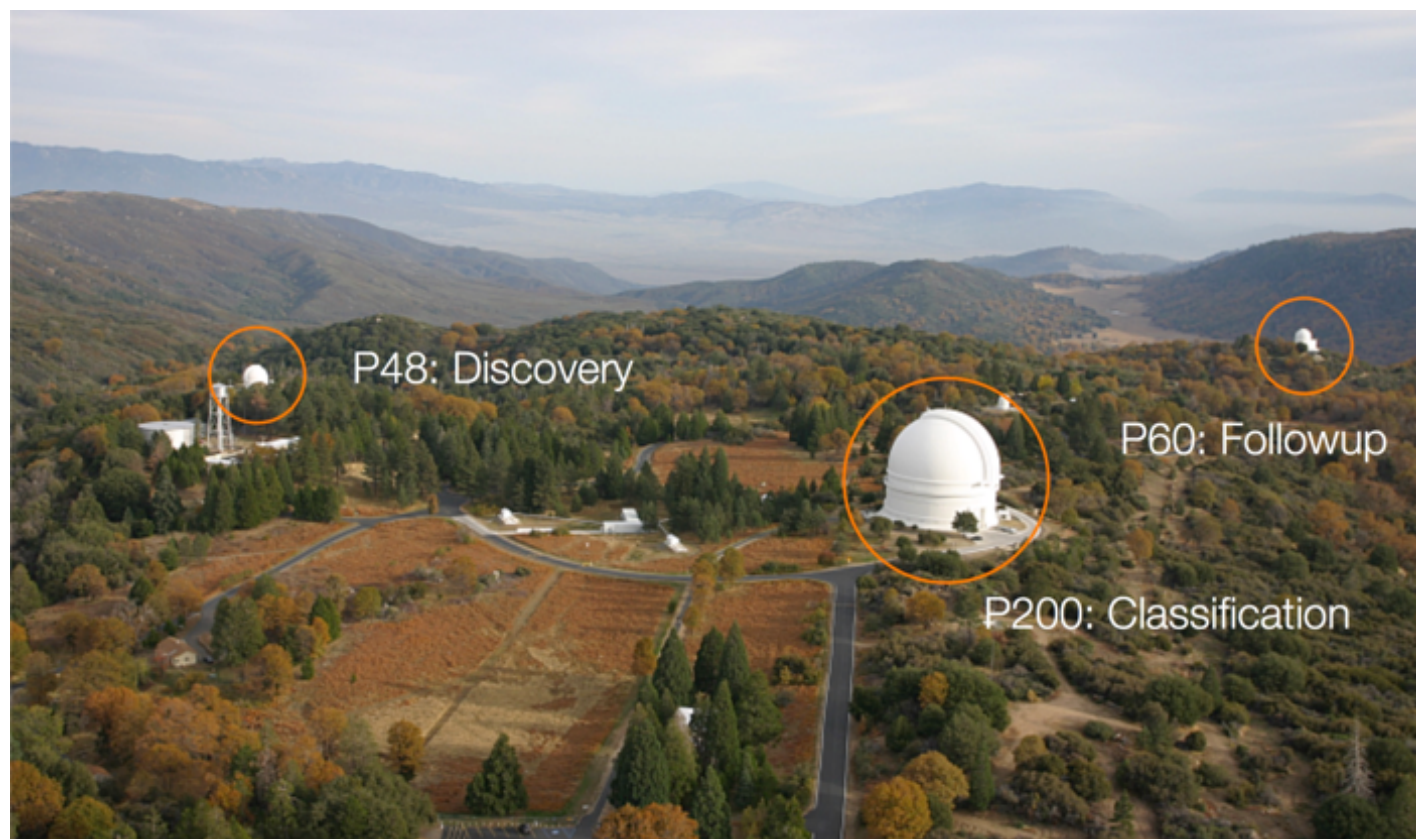
Michael W. Coughlin
University of Minnesota

On behalf of many folks

January 14, 2022



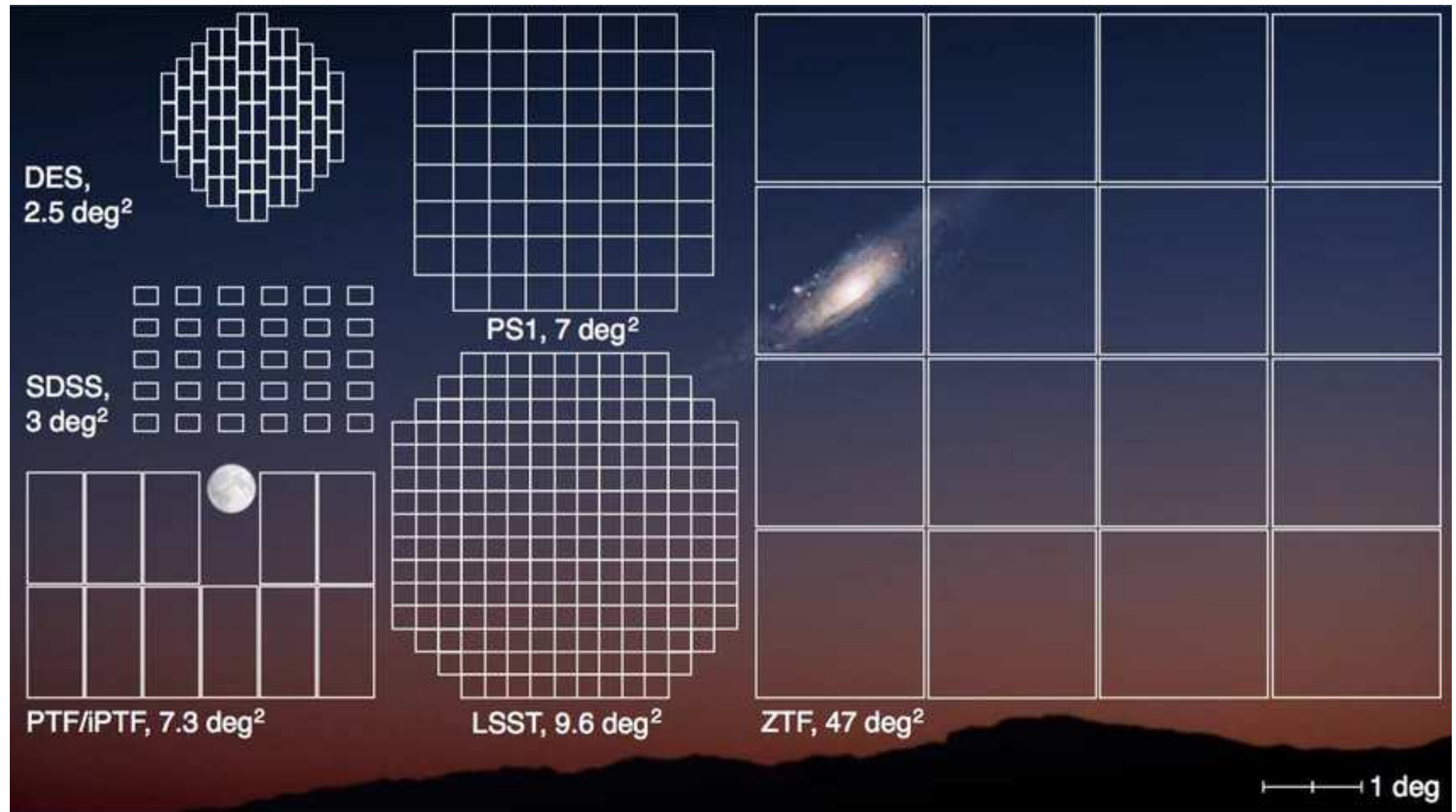
The Palomar Time Domain Astronomy System



And soon, a second SED Machine on the Kitt Peak 2.1 meter



Landscape of Optical TDA



- The night sky is imaged at 17.5 mag by ASAS-SN (both hemispheres)
- The northern sky is covered by ATLAS, ZTF, and PS-1 to 19, 20.5, 21.5 over roughly two nights (ZTF issues real time, data-rich alerts)
- BlackGEM (21-22 mag; Chile) will start routine operation within this year
- Rubin is expected to become operational in a few years

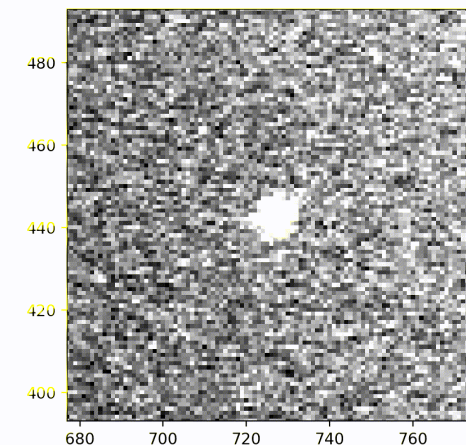
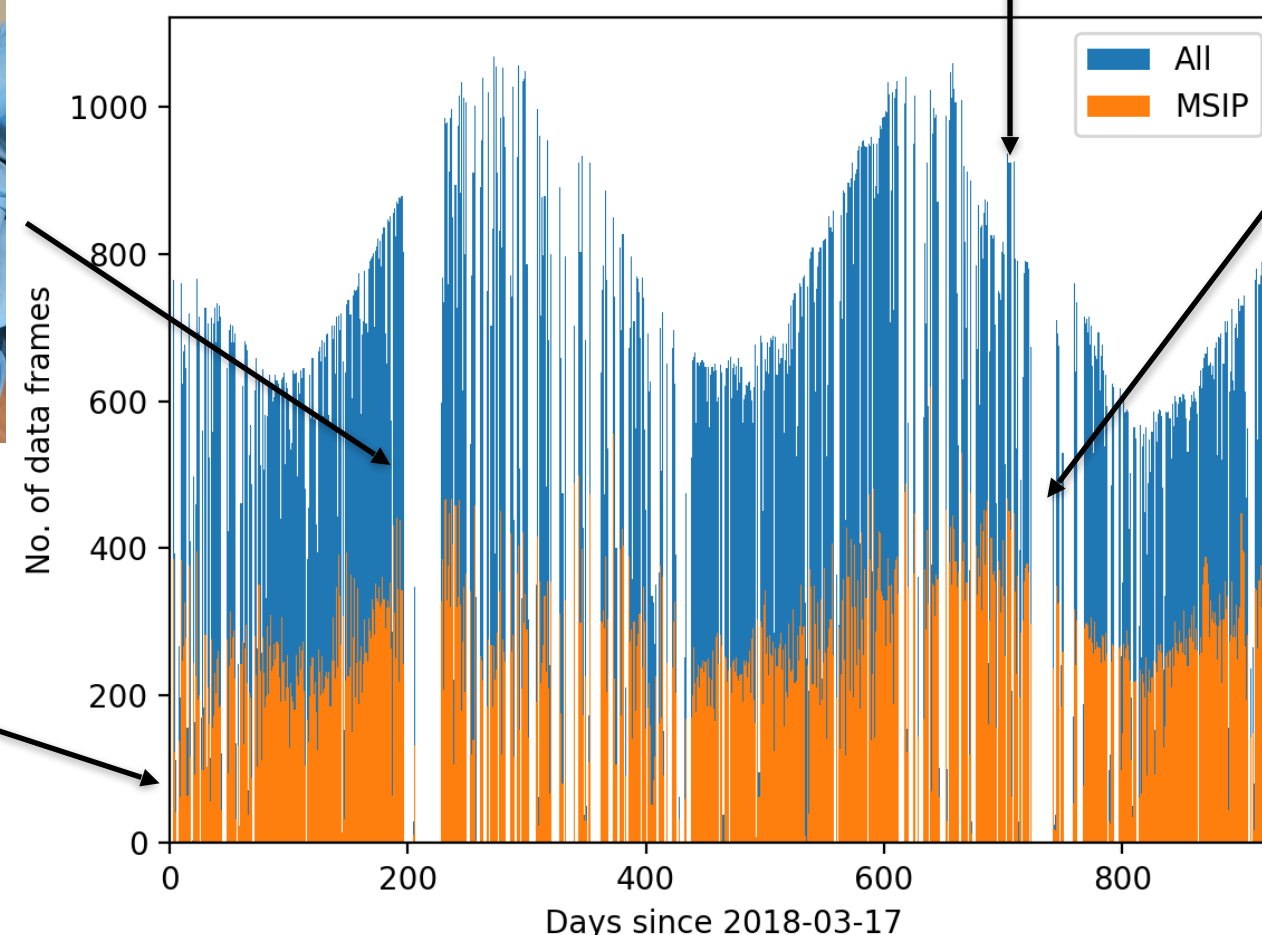


Observing Nights

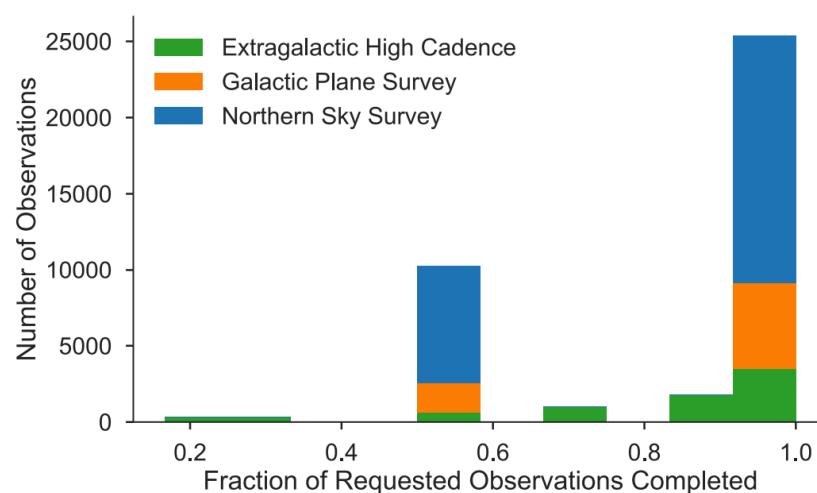
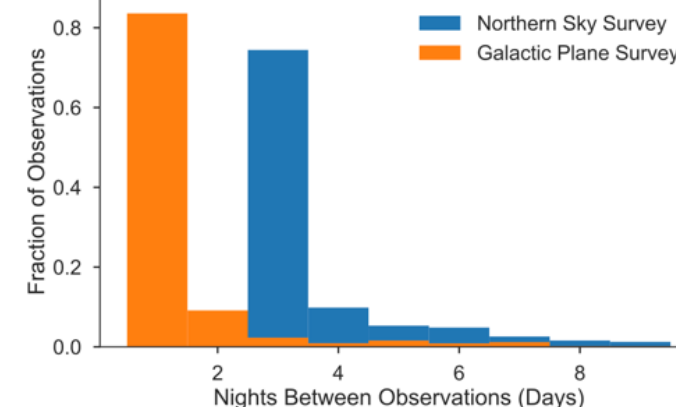


Cryocooler leak

Rained first 2 nights



Wettest March in 25 years



- The most complex time-domain survey ever scheduled
 - ZTF-I: 40% MSIP, 40% partnership, 20% Caltech:
 - Cadenced surveys, time constrained, deep drilling, ...
 - 88+ survey configurations over Phase I
 - >3 changes / month
 - 4137 filter exchanges
 - 3121 ToOs
- If you have questions, ask the next speaker ;)



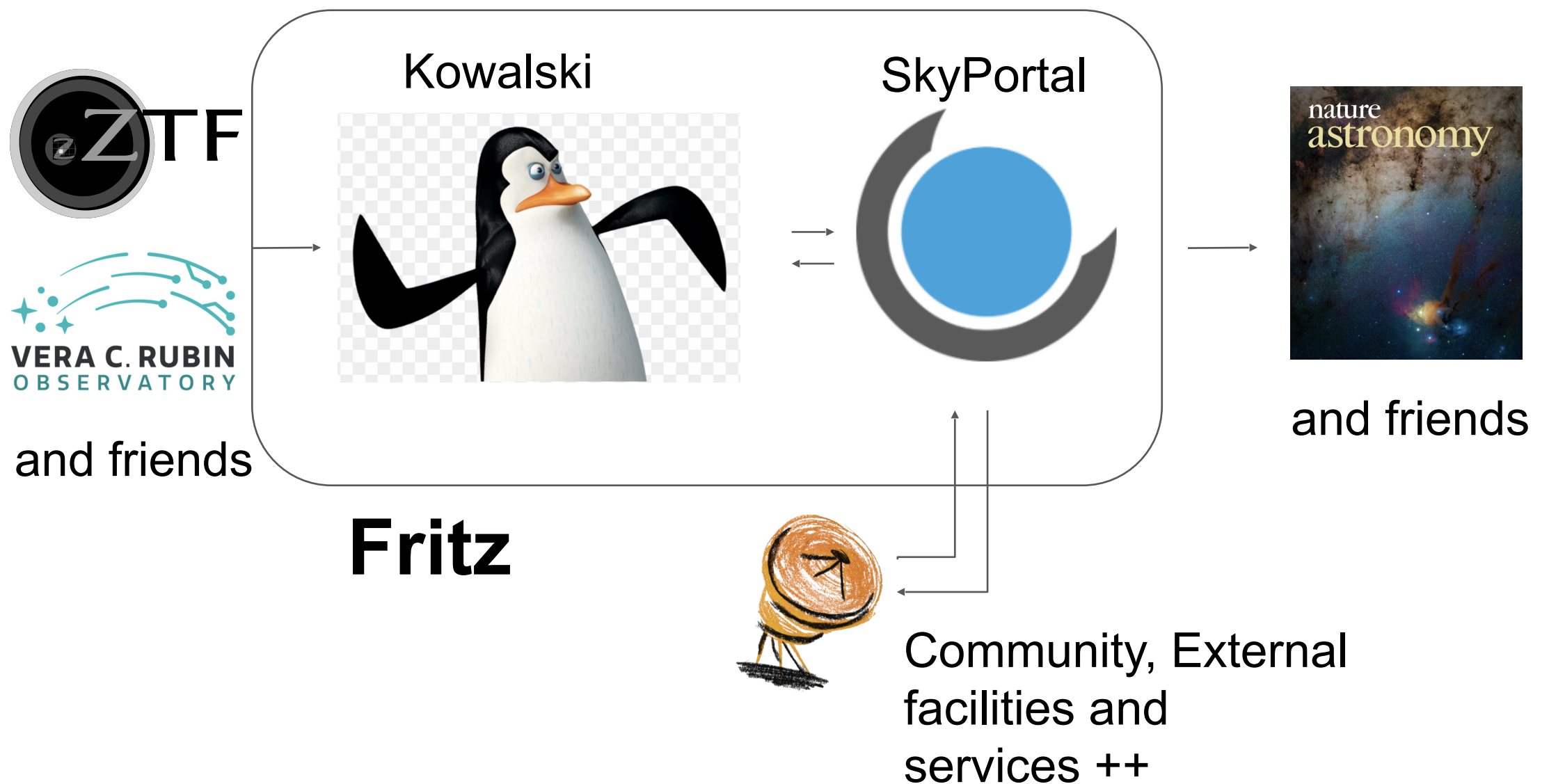
ZTF = 1/10 LSST

		
No. of sources	1 billion	37 billion
No. of detections	1 trillion	37 trillion
Annual visits per source	1000 (2+1 filters)	100 (6 filters)
No. of pixels	600 million (1320 cm ² CCDs)	3.2 billion (3200 cm ² CCDs)
Field of view	47 deg ²	9 deg ²
Hourly survey rate	3750 deg ²	1000 deg ²
Nightly alert rate	1 million	10 million
Nightly data rate	1.4 TB	15 TB



Fritz: Schematic overview

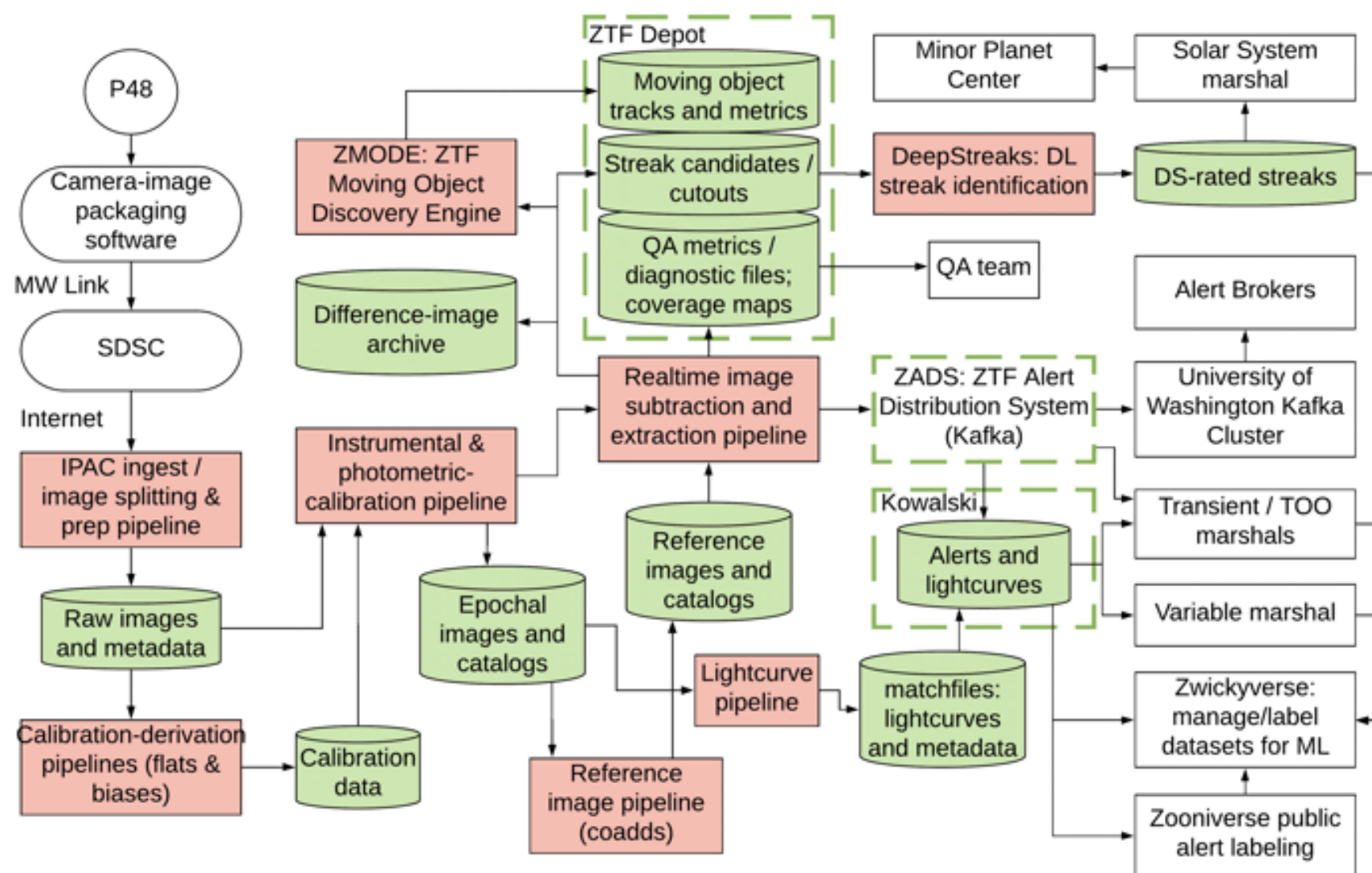
Observe → Mine/Discover → Study and Characterize → Profit!



<https://docs.fritz.science/>



ZTF-I: data/processing flow



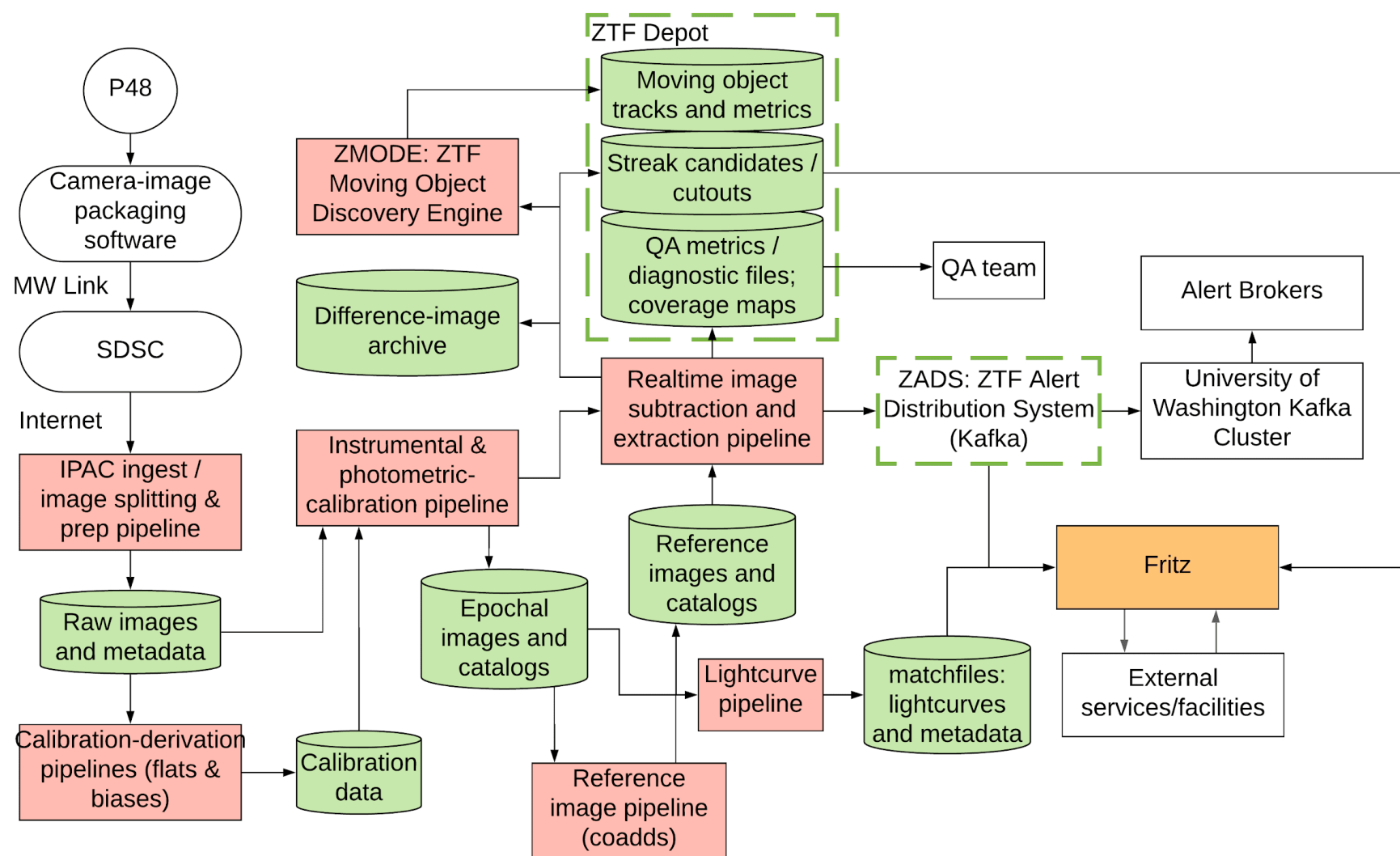
Single night	8h40m
# on-sky exposures	~700
Raw image data	~1 TB
Real-time data products	~4 TB
# unvetted 5σ alerts	~10 ⁵ - 10 ⁶
# ML-vetted alerts	~10 ³ - 10 ⁵
# unvetted streaks	~10 ⁴ - 10 ⁶
# ML-vetted streaks	~10 ² - 10 ³

Nominal survey	3 x 260 n
Volume of data products	~3.2 PB
Volume of ref images	~60 TB
# CCD quad ref images	~2.8 x 10 ⁵
Volume of matchfiles	~50 TB
# matchfiles	~2.8 x 10 ⁵
# single-epoch PSF-fit source measurements	~800 B
# single-epoch aperture source measurements	~230 B

See Masci+ 2019



ZTF-II: data/processing flow



Single night	8h40m
# on-sky exposures	~700
Raw image data	~1 TB
Real-time data products	~4 TB
# unvetted 5σ alerts	$\sim 10^5 - 10^6$
# ML-vetted alerts	$\sim 10^3 - 10^5$
# unvetted streaks	$\sim 10^4 - 10^6$
# ML-vetted streaks	$\sim 10^2 - 10^3$

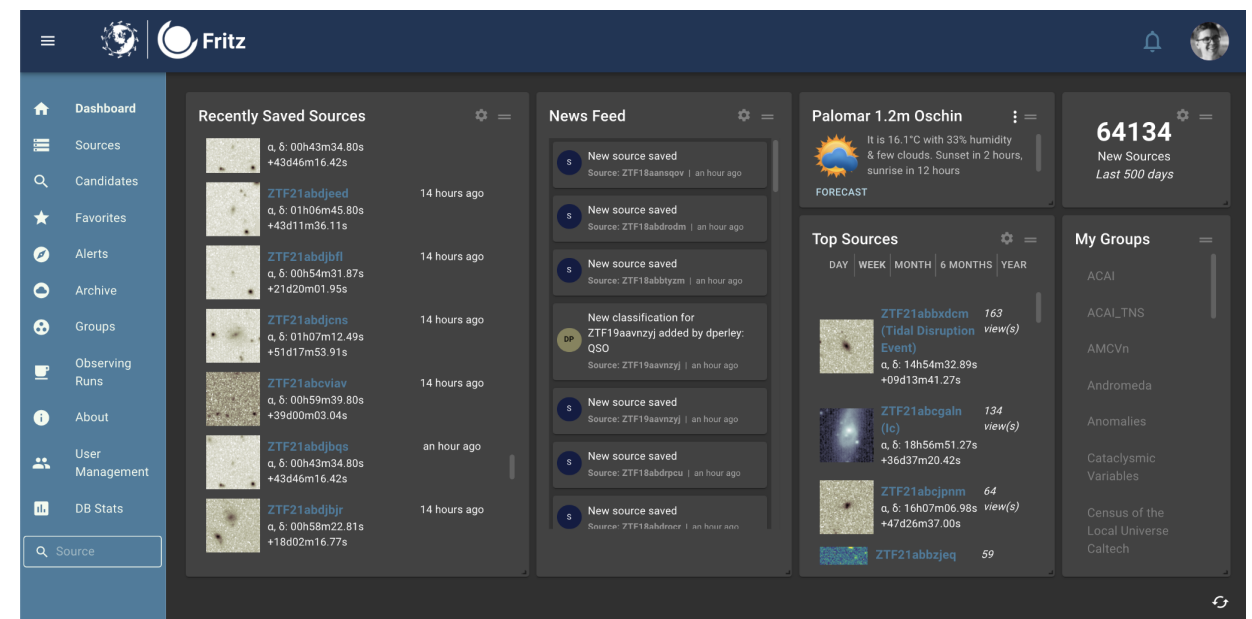
Nominal survey	3 x 260 n
Volume of data products	~3.2 PB
Volume of ref images	~60 TB
# CCD quad ref images	$\sim 2.8 \times 10^5$
Volume of matchfiles	~50 TB
# matchfiles	$\sim 2.8 \times 10^5$
# single-epoch PSF-fit source measurements	~800 B
# single-epoch aperture source measurements	~230 B



Fritz: science data platform for ZTF-II

- Scalable, API-first system, with fine-grained access control
- Multi-survey data archive and alert broker
- Interactive, mobile-friendly collaborative platform for transient, variable, and Solar system science cases
- Workhorse for ML applications: classification and labeling at scale
- Follow-up observation management: robotic and classical facilities

Initiated in Feb 2020
Beta up in Sep 2020
MVP live in Nov 2020



<https://github.com/fritz-marshall/fritz>

<https://docs.fritz.science>



Fritz: features

- Data archive
 - ZTF alert stream
 - ZTF light curves



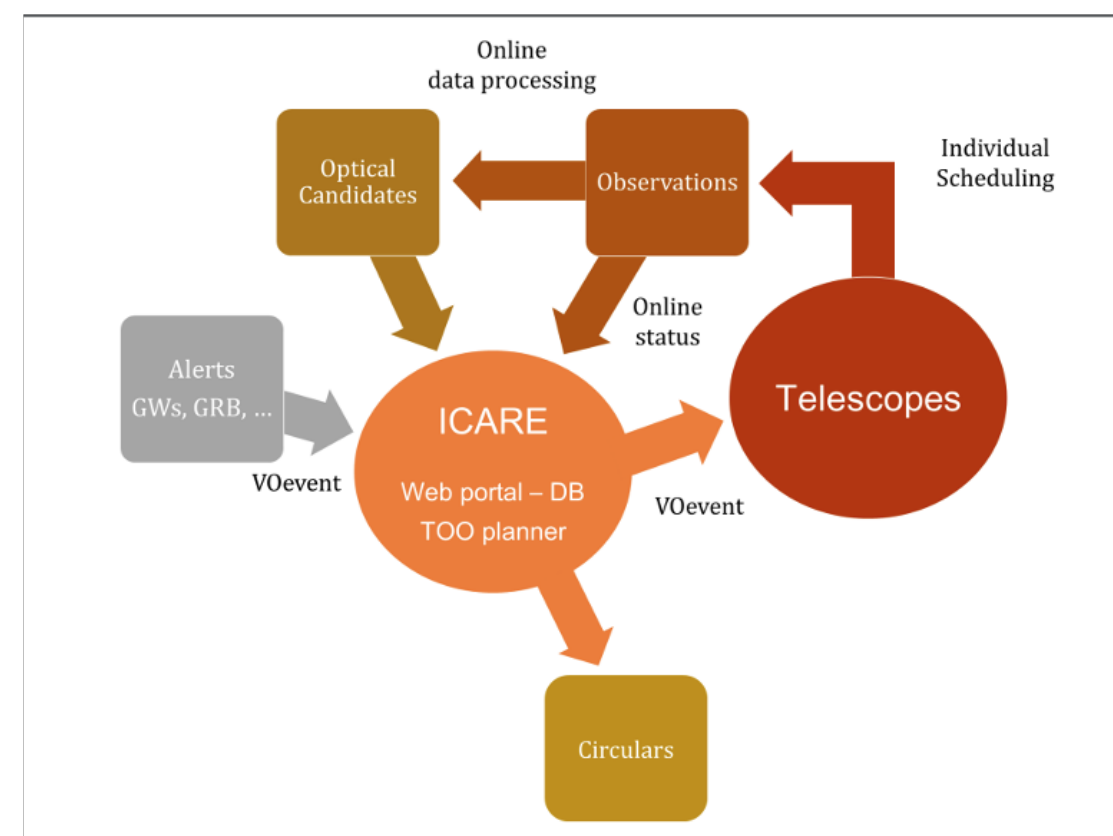
MMA Development

Inherits from two very successful projects during O3 and beyond:
GRANDMA's iCARE and GROWTH's ToO Marshal

Events

☒ AMON ☒ AllSky ☒ BBH ☒ BNS ☒ Fermi ☒ GRB ☒ GW ☒ HighMass ☒ IMBH
☒ LVC ☒ MDC ☒ MassGap ☒ NSBH ☒ Terrestrial ☒ long ☒ retracted ☒ short
☒ transient

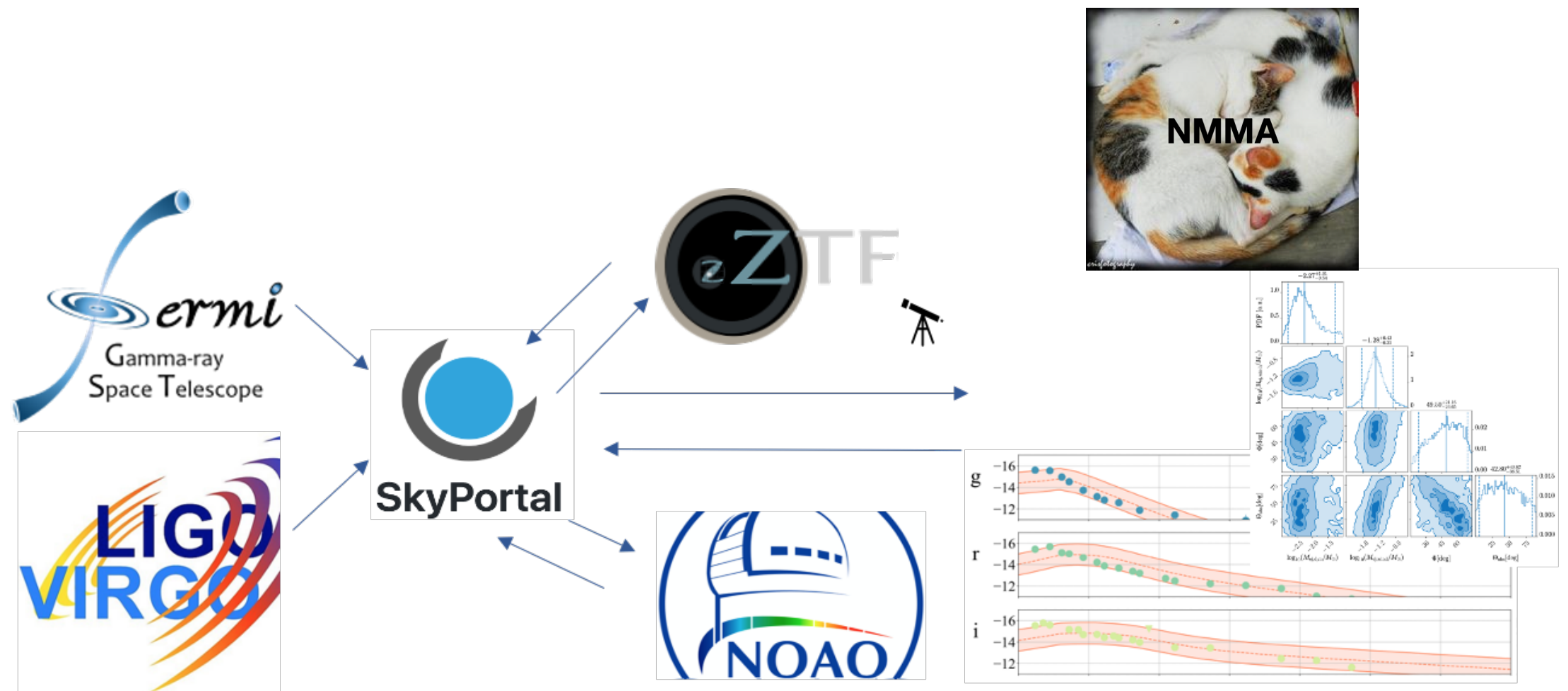
210511 20:51:24 (18 hours ago)	Fermi GRB
210511 11:26:39 (a day ago)	Fermi long GRB
210511 10:58:57 (a day ago)	Fermi long GRB
210510 19:20:01 (a day ago)	Fermi short GRB
210509 13:53:40 (3 days ago)	Fermi GRB
210509 12:15:04 (3 days ago)	Fermi short transient
210508 22:27:35 (3 days ago)	Fermi GRB
210508 21:37:25 (3 days ago)	Fermi short transient
210508 21:30:44 (3 days ago)	Fermi short transient
210507 19:01:51 (4 days ago)	Fermi GRB



Ahumada et al. 2105.05067, Anand and Coughlin et al.
2009.07210, Andreoni and Goldstein et al. 1910.13409, Antier et al:
1910.11261, 2004.04277, Coughlin et al.: 1907.12645, etc.



A vision for the O4 workflow



Goal: Inform follow-up decisions

Goal: Remove difference between code we use to vet candidates in low latency and code we use to do science



New features in development

- MOC-healpix Postgres 14 based object / galaxy cross-matches with skymaps: <https://github.com/skyportal/healpix-alchemy>
- Existing ACL's for telescope triggering interfaces / programs
- Observation Planning page with network level scheduling
- One stop shop for light curves (and therefore light curve fitting suite) and telescope limits (and therefore parameter constraints)



Sarah Antier, OCA



Leo Singer, Goddard

Epic tracking MMA work:

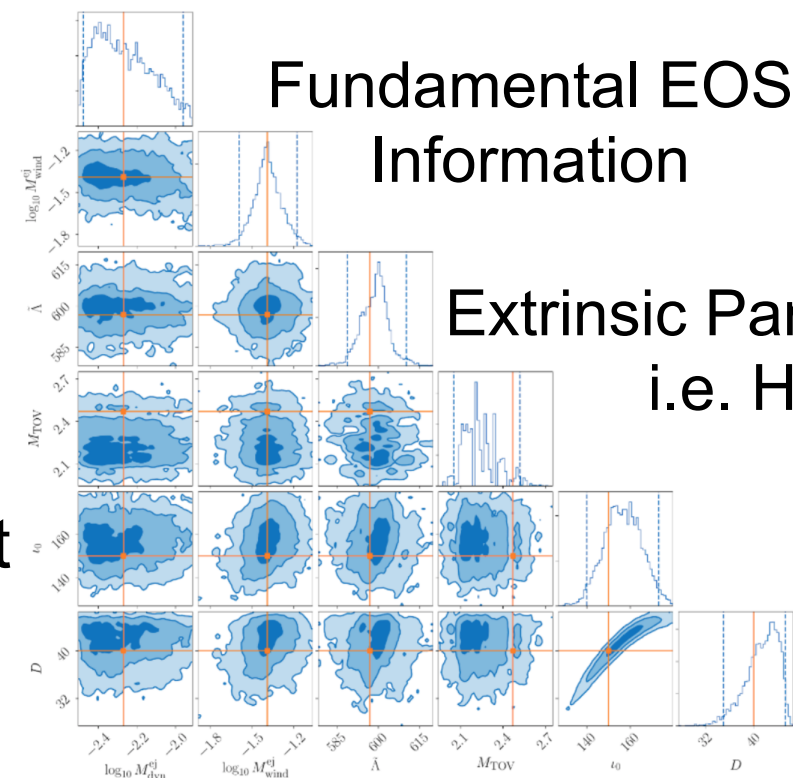
<https://github.com/skyportal/skyportal/issues/2051>

Hope to finish by summer and stress test active neutrino and GRB programs

Ejecta Parameters

Fundamental EOS Information

Extrinsic Parameters
i.e. H_0





Lessons Learned and Conclusions

open source is

- Fritz/SkyPortal/Kowalski is open source
 - A huge part of devs is volunteer labor of love
 - While the core dev team is relatively small (in terms of effective person-hours), dozens of people have [contributed](#) meaningfully
- Leveraging what GitHub/OSS has to offer
 - Issues to track bugs and feature requests
 - PRs + thorough code review
 - GitHub Actions as the CI/CD platform
 - Don't argue about style, enforce pre-commit hook (black, flake8, eslint...)
- Project management tool: ZenHub
 - Should be as close as possible to GH

Testing!

- You are not testing your code enough
 - From unit testing to integration testing through API and frontend, every bit is essential
 - Is that docker image still building from scratch? Note the word "continuous" in CI
 - Never underestimate the scale of a disaster that six innocently-looking lines of code can bring
 - Database migrations should be tested both ways - roll-backs are more common than we'd wish
 - Understand (and embrace!) flakiness
- Staging environment
 - Helps catch a lot of bugs before they have a chance to reach production, e.g. innocently-looking migrations that can take forever

Team and community is

- Extraordinary individuals with a broad range of expertise
 - Critical but fair and open-minded review of ideas and code allows to iterate fast, converging on better solutions
 - Staying in sync: Slack + Daily 15-min stand-ups + weekly 1h meetings
- Constructive community feedback is essential for success
 - Enormously useful for finding/fixing bugs and implementing new features
 - Need to be clear about the communication channels: a dedicated Slack channel for smaller issues + GH issue templates for larger stuff
 - Critically important: prioritization + clear big picture for the project

Production is really hard!

- Even harder is to deliver updates/new features to prod
- Testing *is* your friend, but it won't catch everything that can happen
 - Running a subset of the test suite on a read-only replica of the prod db
- Weekly (at least) deployment to prod
- Resilient infrastructure for deployment
 - Monitoring the performance of the different components
 - Query Insights on the GCP
 - API endpoint response times, temporal evolution



Thank you!



Fritz: features

- Open source (free to use, modify, and distribute)
- API-first system: rich APIs for machine usage
- Powerful alert stream enhancement/filtering capabilities
- Extendible & scalable design: async Python backends, React/Redux frontend
- Fine-grained access control
- Authentication via OAuth
- Real-time Slack-like messaging, notifications
- Rich visualization capabilities
- Follow-up management
- Distributed computation via Dask
- Docker compose or Kubernetes deployment
- Well-tested, extensive docs, CI/CD

See the [tutorial](#)
presented at the
[LSST Brokers](#)
[workshop](#)

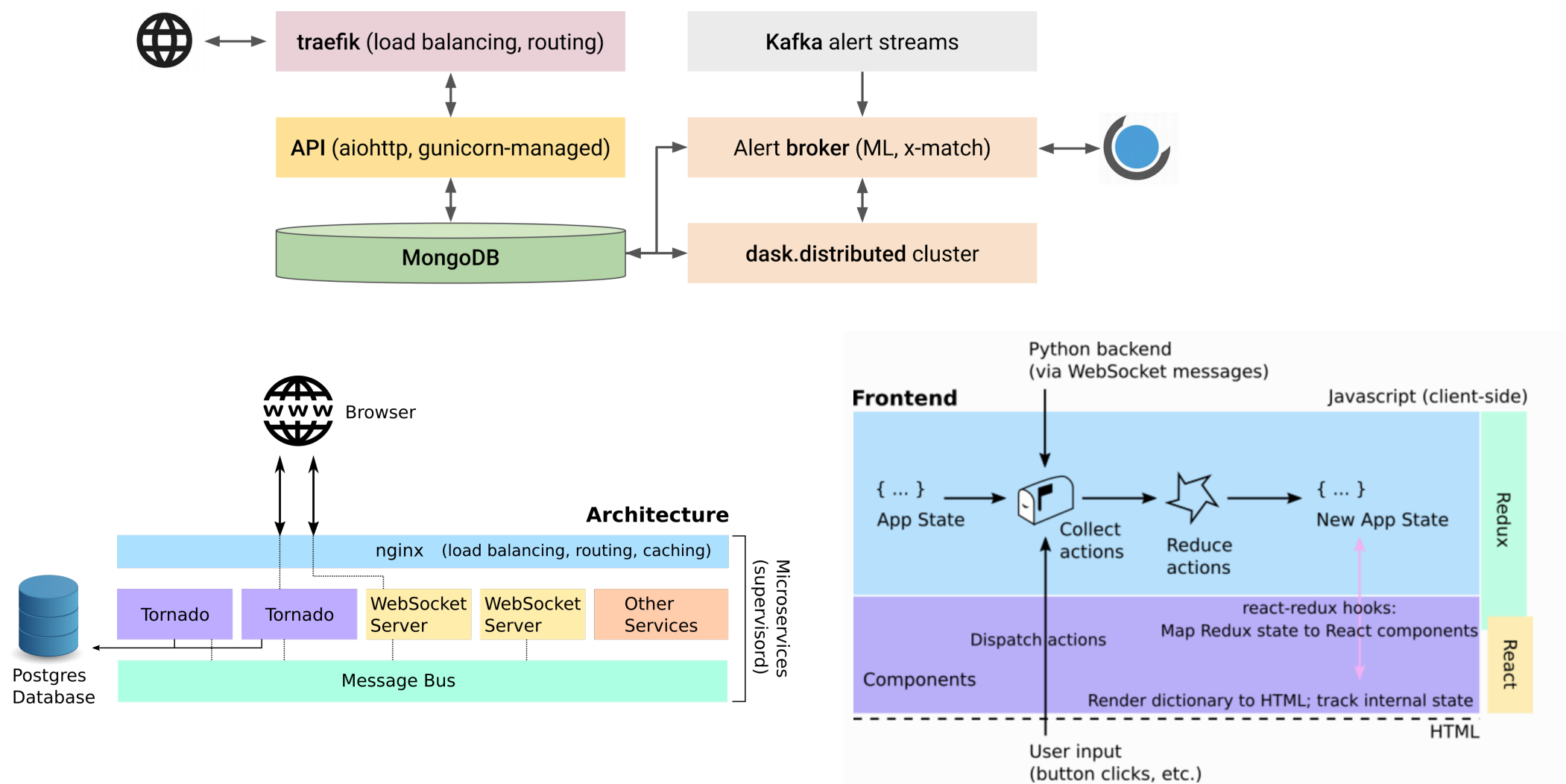
+

[https://](https://docs.fritz.science/)
docs.fritz.science/



Fritz: features

Extendible & scalable design: async Python backends, React/Redux frontend





Fritz: features

Alert stream enhancement & filtering capabilities

- MongoDB aggregation pipelines
- Performant for filters of different complexity
- Simple `dask.distributed`-based service(s)
- Public alert DBs for filter design/debugging
- Multiple ML models (ACAI)
- Cross-matches
 - ~10 external catalogs
- Filtering enhanced data
- Watchlists
- Automated checks, no filter code audit
- [Can post results from external sources]

