

AMPEL deployment at DESY

Operation and performance

Jakob van Santen <jakob.van.santen@desy.de>
AMPEL Workshop, DESY Zeuthen, 2019-05-23



AMPEL framework

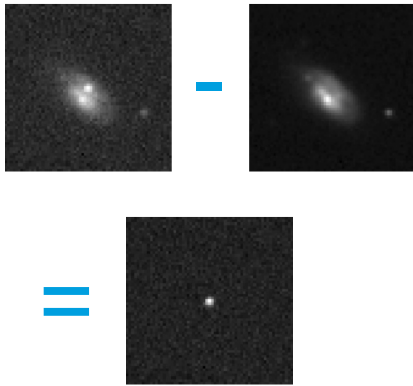
Alert Management, Photometry, and Evaluation of Light curves

- AMPEL is a streaming data analysis framework with a focus on astronomy, currently in active development
- ~12k lines of Python3
- Core does things that astronomers are bad at:
 - Provenance tracking
 - Data and computation de-duplication
 - Horizontal scaling
- Contributed plug-ins do things that astronomers are good at:
 - Account for finicky details of individual instruments
 - Calculate object features
 - Analyze population properties



Transient-astronomical data

Photometric point (data point)

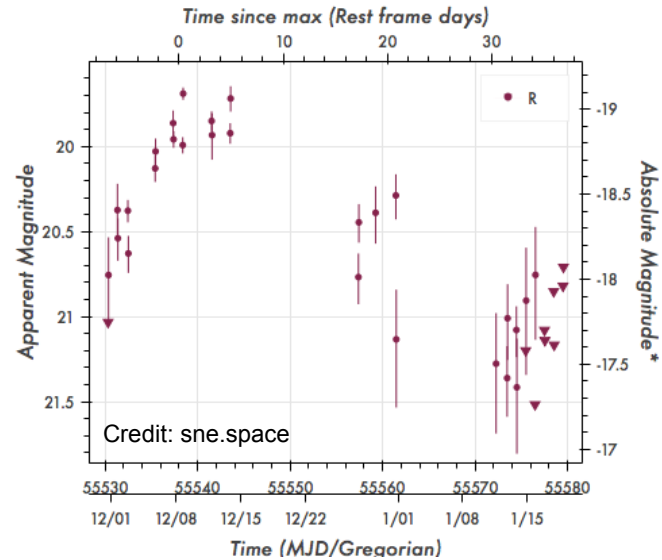


```
{  
  "direction"  
  "time"  
  "intensity"  
  "metadata"  
  "_id" : 674294034215015004,  
  "ra" : 31.7952798,  
  "dec" : -9.0246254,  
  "jd" : 2458428.7940394,  
  "magpsf" : 20.2646770477295,  
  "exptime" : 30.0,  
  "fid" : 1,  
  "field" : 349,  
  ...  
}
```

Tasks:

- Accept/reject based on data quality, location, etc

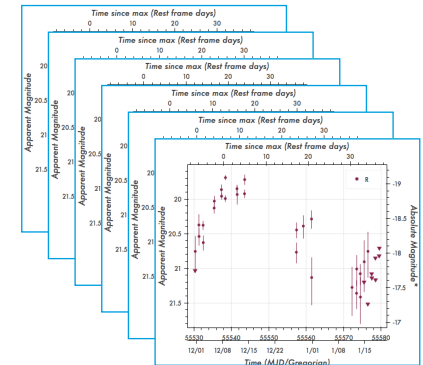
Light curve (time series)



Tasks:

- Update object features

Population (panel)



Tasks:

- Estimate properties of underlying population
- Rank by importance, **request follow-up paintings from other telescopes**
- ...

ZTF Kafka stream

Topic per night and observing program
(16 partitions each)

```
ztf_20181115_programid1  
ztf_20181115_programid2  
ztf_20181116_programid1  
ztf_20181116_programid2  
ztf_20181117_programid1  
ztf_20181117_programid2
```

upstream (UW) retains messages for ~1 week

Message per detection
(Avro payload with single record)

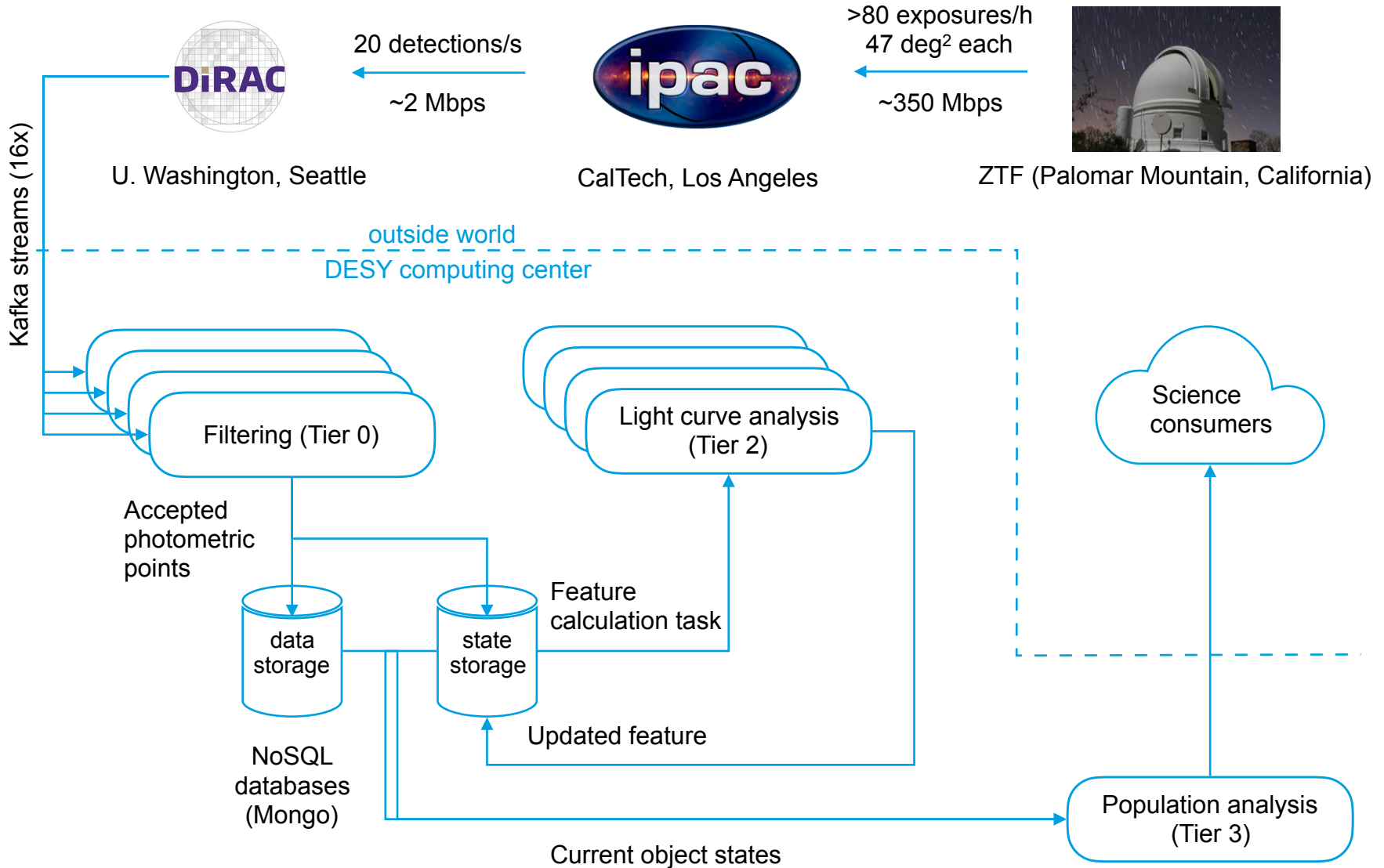
```
[ 'schemavsn',  
  'publisher',  
  'objectId',  
  'candid',  
  'candidate',  
  'prv_candidates',  
  'cutoutScience',  
  'cutoutTemplate',  
  'cutoutDifference' ]
```

internal object name
id of last detection
last detection
previous detections
image data

typically 60-100 kB

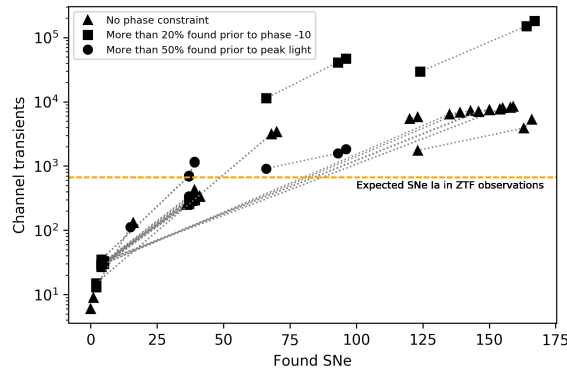
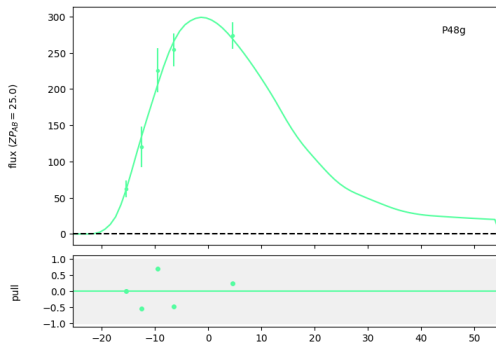
Live AMPEL instance at DESY

Real-time analysis of Zwicky Transient Facility data



Current applications

Searching for associations between supernovae and high-energy neutrinos



Optimizing transient selections

Monitoring the sky for gravitationally-lensed supernovae

#ztf_lensing

☆ | 👤 7 | 🗨️ 0 | ⚙️ Add a topic

Tuesday, November 27th

AMPEL-live APP 4:01 PM

UPDATE! Alert summary for 2018-11-27.
 LONE WOLF! It seems that there was just one transient which passed the filters.
 Guess it's better than nothing... 🐺
 The results are summarised below. That one result is shown below.

AMPEL APP 4:01 PM
 Summary: 2018-11-27

```
1 | , ztf_name, ra, dec, magpsf, sgscore1, rb, most_recent_detection, fir
st_detection, n_detections, distnr, distpsnr1, isdiffpos, _id, T2-
snscosmo_info_success, T2-snscosmo_info_chisq, T2-
```

🌳 **TRANSIENT NAME SERVER**
SEARCH
BOTS
LIGO GW

LOGIN
✉️
?

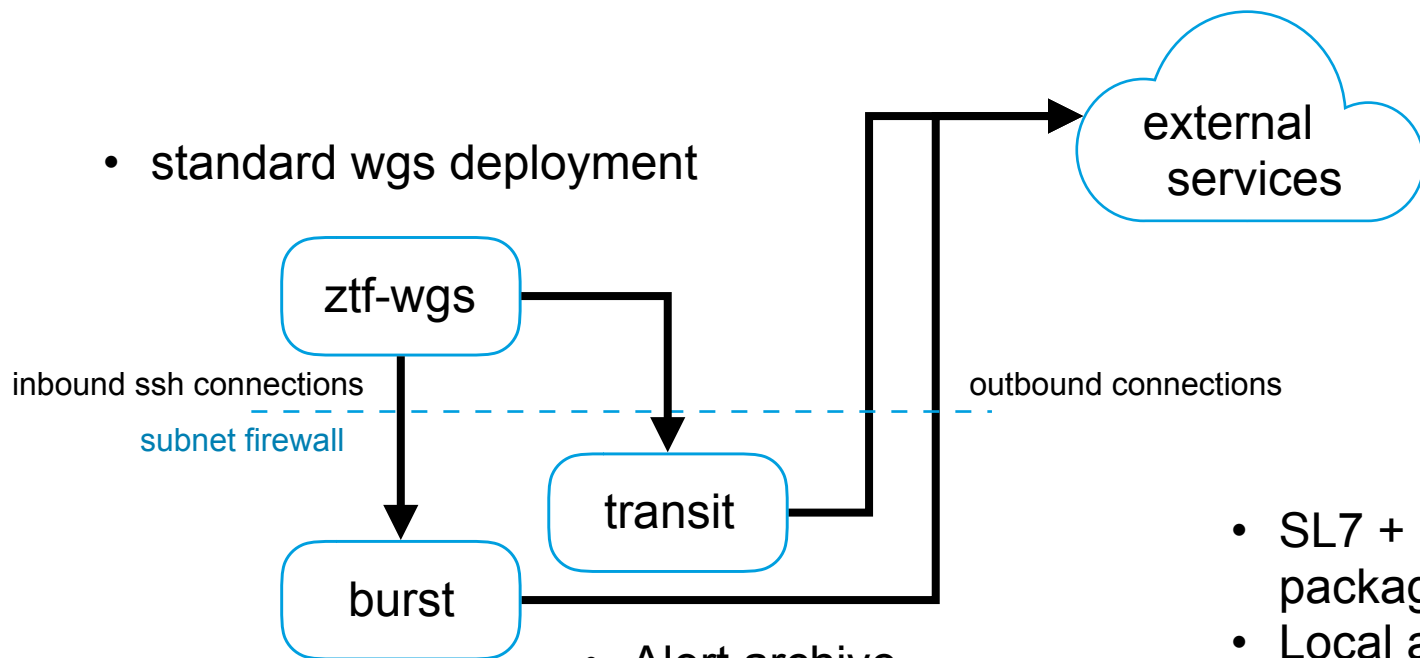
Selecting transients for real-time public distribution

ID	Name	Reps	Class	RA	DEC	Type	Discovering Group/s	Disc. Internal Name	Public	Discovery Mag	Discovery Filter	Discovery Date (UT)	Sender
31803	AT_2018jjs	1		05:08:03.839	-03:32:58.32	ZTF	ZTF18acrvakb	ZTF18acrvakb	Y	19.1079	r-ZTF	2018-11-28 09:00:29	ZTF_AMPEL_MSIP
31802	AT_2018jjr	1		05:41:26.957	-01:54:09.44	ZTF	ZTF18acrnanad	ZTF18acrnanad	Y	18.0792	r-ZTF	2018-11-28 08:34:03	ZTF_AMPEL_MSIP
31801	AT_2018jjq	1		05:39:25.178	-06:09:00.03	ZTF	ZTF18acrukpd	ZTF18acrukpd	Y	17.8615	r-ZTF	2018-11-28 08:34:03	ZTF_AMPEL_MSIP
31800	AT_2018jjp	1		04:31:45.488	+79:27:01.76	ZTF	ZTF18aclmvt	ZTF18aclmvt	Y	18.4496	r-ZTF	2018-11-28 08:47:57	ZTF_AMPEL_MSIP

Operations

Lightweight deployment

- standard wgs deployment

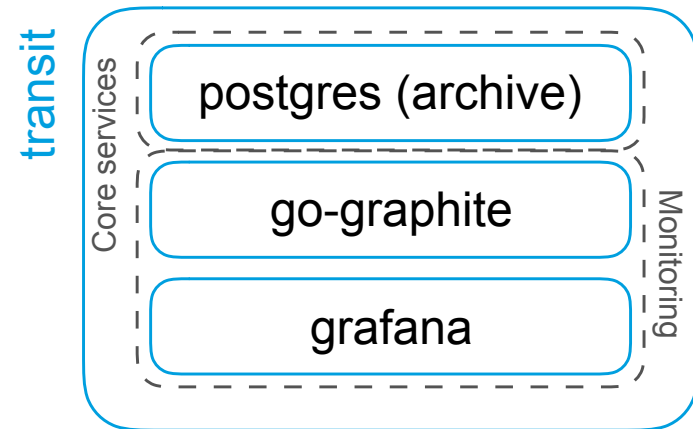
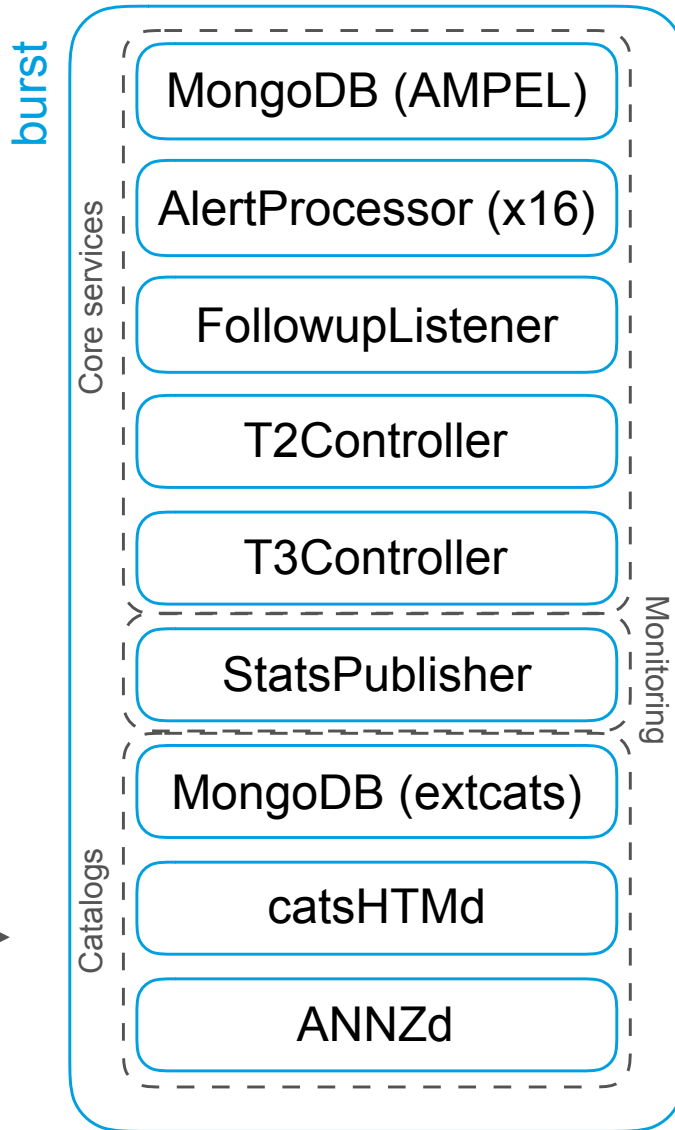


- Live MongoDB
- Catalogs
- Live stream processing

- Alert archive
- Offline analysis

- SL7 + minimal package set
- Local accounts, isolated from DESY services (no AFS, no Lustre)
- DV manages firewall and storage, JvS the rest

AMPEL stack



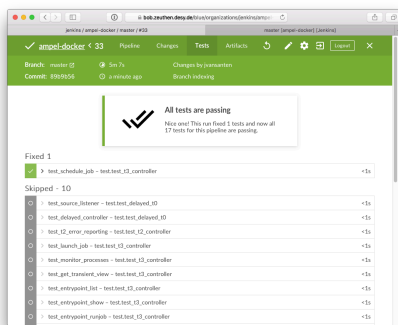
See M. Giomi's talk →

- Services distributed as Singularity containers
- Orchestrated with `singularity-stack` (custom `docker swarm` workalike)

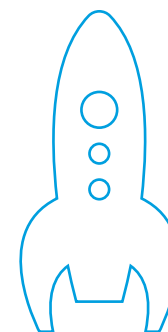
Container workflow



- ## Assemble Dockerfile
- Dependencies (conda)
 - Tagged version of core Ampel projects and contributed plug-ins (git submodules)



- ## Build with Jenkins
- Build Docker image
 - Run unit & integration tests
 - Convert to Singularity format
 - Publish Singularity image to DESYcloud



- ## Deploy
- Download tagged Singularity image
 - Record stack configuration (image tags, bind mounts, env vars, etc.)
 - Deploy with singularity-stack

~20 minutes start to finish

Performance

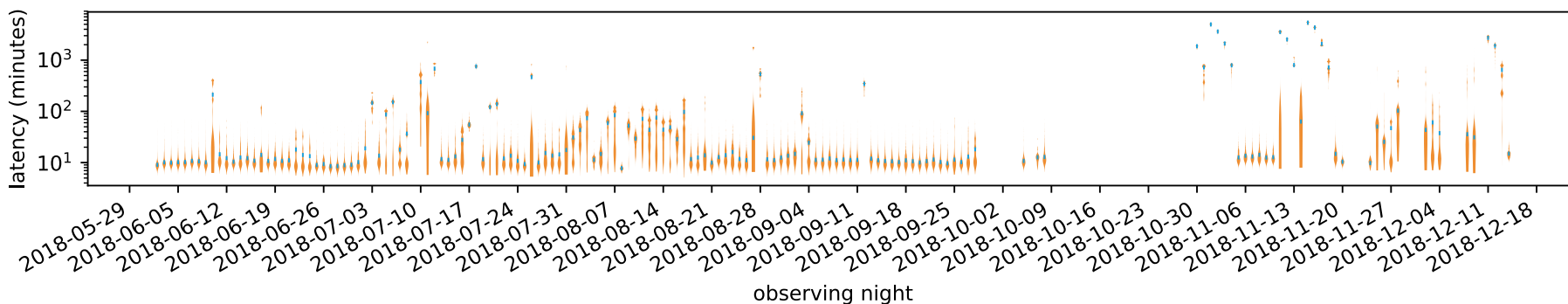
Data volumes

- ZTF alert distribution began in June 2018
- Performance as of 1.1.2019:
 - 64M alerts (one Kafka message each)
 - 4.9 TB (2.9 TB de-duplicated)
 - Latency from shutter close at Palomar to alert ingestion at DESY typically < 10 min

Stable running

Camera repair

Winter rains



Performance limits

- All AMPEL tiers are embarrassingly parallel and communicate over sockets
 - Scale horizontally
 - Scale to multiple nodes/sites (with better orchestration)
- Central MongoDB is a potential bottleneck (modulo sharding)
- Single Mongo instance on `burst` can sustain 1250 alerts/s (~3x LSST rate) with current filter channels (1-2% passing rate)

Summary

- AMPEL has been processing ZTF alerts at DESY for nearly one year. Current applications include:
 - TDE detection with ZTF (see talk by R. Stein)
 - Early SNIa candidate detection (see talk by V. Brinnet)
 - Neutrino-optical transient coincidence search (see talk by L. Rauch)
- AMPEL is modular (instrument/science plug-ins) and its releases containerized for portability and reproducibility.
- The processing pipeline scales horizontally; LSST-scale alert rates can be handled easily.