

# Automating ‘Inbox Zero’:

An approach to processing system-generated messages

OPFCON  
June 9, 2020

David Cirella  
Digital Preservation Librarian, Yale University Library

---

# System Messages

- Preservation System messages via email
    - Success & error messages
    - Various system processes including: ingest, access, integrity checking
  - Large volume: 529,258 messages (2017 - 2020)
-

# Handling Scale

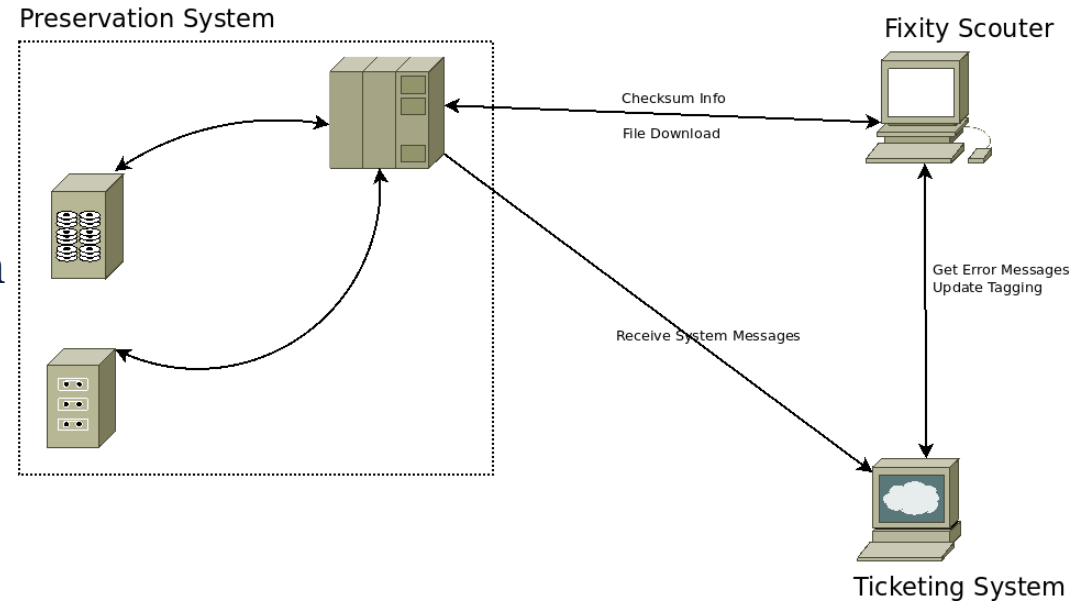
- Partially solved with ticketing system
    - Tagging based on message content
    - Routing and forwarding
  - Not a complete solution for messages that require action
-

# Storage Integrity Messages

- Nightly scheduled fixity check of files
  - Messages to report on failure for various reasons
    - Compute checksum
    - Access to files
-

# Fixity Scouter

- Python script that leverages APIs from ticketing system and preservation system
  - ▣ Parse each message
  - ▣ Generate calls to get checksum value and hash algorithm from preservation system
  - ▣ Download files
  - ▣ Run checksum to perform local fixity check
  - ▣ Update tags based on result
  - ▣ Data transfer cap/throttling



# Outcomes

- Automated processing of error messages at scale
    - ▣ Verify file integrity
    - ▣ Identify cause of error messages
  - Identify patterns that contribute to ‘error’ messages
    - ▣ Storage recall speed
    - ▣ Outages
-

---

Thank you

---