SMA-X: Versatile Information Sharing

Attila Kovács, Paul Grimes, Christopher Moriarty, and Robert Wilson
Center for Astrophysics | Harvard & Smithsonian

INTRODUCTION

The **SMA eXchange** (SMA-X) is a real-time data storage convention built atop a central **Redis** database, and originally developed for the Submillimeter Array (SMA). It supports flexible storage of hierarchical data, with the possibility of fast, atomic retrieval of specific leaf elements and sub-trees, including associated metadata (types, dimensions, timestamps, and origins, and more).

SMA-X provides fast, efficient, and low-latency data sharing with minimum network traffic and network complexity.

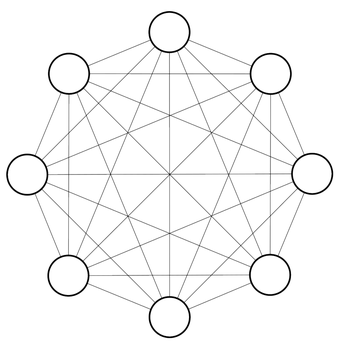

Fig. 1 Decentralized P2P network. Information is shared on $O(N^2)$ links

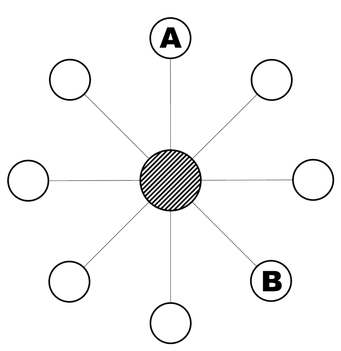

Fig. 2 Centralized network: (a) P2P with central decision making, or (b) public message board with decentralized decision making.

SMA-X is superior to P2P, with its flexibility and transparency, and because to supports decentralized decision making, and monitoring

REPOSITORIES

SMA-X will be open-sourced in Summer 2024

It will be available via the **GitHub** repositories:

`Smithsonian/smax-server`	Server configuration
`Smithsonian/smax-clib`	C/C++ library and tools
`Smithsonian/smax-python`	Python3 library
`Smithsonian/smax-postgres`	PostgreSQL connector application

DATA TYPES

Basic data types

`int[8,16,32,64]`	Integer value(s), with or w/o width specification
`float`	single-precision floating point value(s)
`double`	double-precision floating-point value(s)
`boolean`	string(s), with JSON-style escaping
`string`	logical true/false value(s)
`struct`	reference to substructure content

Hierarchical data

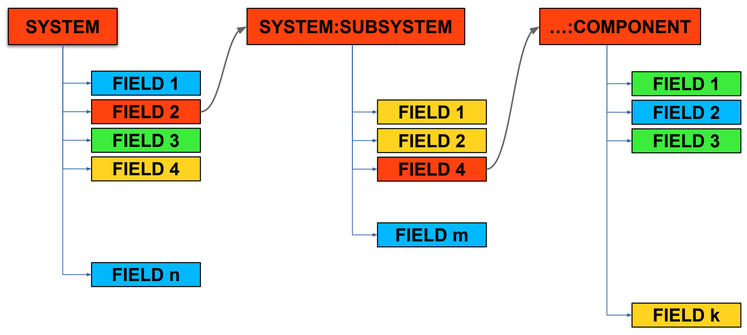

Fig. 3 Arbitrary data hierarchies are stored as a set of linked Redis hash tables. Each table contains leaf nodes and sub-structure references

Associated Metadata

All branches and leaf nodes have associated metadata that describe the value(s) stored

`<types>`	Published data type (see above)
`<dimensions>`	Array dimensions, e.g. “4 5 8” for a 4x5x8 array
`<origins>`	host and program that published the value
`<timestamps>`	Precise UNIX time data was published
`<reads>`	number of times data was accessed
`<writes>`	serial number (number of times updated)

PUB/SUB update notifications

Each time data (leaf-nodes) or entire branches, are updated in SMA-X, a set of notifications are sent immediately, on dedicated channels for each modified leaf-node and all parent structures. It allows programs to be notified immediately when select variables or structures are updated.

FEATURES

- **High-throughput (~1 million transactions/sec)**
- **Low latency (network-limited)**
- **Atomic transactions**
- **Pipelined (bulk data) access mode**
- **Lazy access (background caching)**
- **Update triggers (semaphores & callbacks)**
- **Automatic widening & narrowing conversions**
- **Resiliency**

Historical Record

We provide a configurable application for creating a detailed historical record in a **PostgreSQL** database (with or without a **TimescaleDB** extension). Both incremental updates and full snapshots are regular intervals are supported. Temporal evolution can be easily plotted and monitored e.g. with **Grafana**.

ADOPTION

SMA-X is currently being used at:

- The Submillimeter Array (SMA)
- MIT Haystack 37-m Telescope Control System (TCS)

CONCLUSIONS

SMA-X can enhance data sharing in telescopes. It reduces network complexity, eliminates unnecessary latencies with minimal network traffic. It supports decentralized decision making, and simplifies monitoring and troubleshooting.

CONTACT

For information, or early access to the software please contact attila.kovacs@cfa.harvard.edu.

The Submillimeter Array is a joint project between the Smithsonian Astrophysical Observatory and the Academia Sinica Institute of Astronomy and Astrophysics and is funded by the Smithsonian Institution and the Academia Sinica.