

A High-performance Tiered Storage System for a Global Spectrum Observatory Network

Ryan Attard¹, Juha Kalliovaara², Tanim Taher¹, Jesse Taylor¹, Jarkko Paavola³, Reijo Ekman³, and Dennis Roberson¹

¹Illinois Institute of Technology, Chicago, USA

²University of Turku, Turku, Finland

³Turku University of Applied Sciences, Finland

E-mails: rattard@hawk.iit.edu, juha.kalliovaara@utu.fi, tahetan@hawk.iit.edu, jtaylor11@hawk.iit.edu, jarkko.paavola@turkuamk.fi, reijo.ekman@turkuamk.fi, robersond@iit.edu

Abstract— This paper describes the RF measurement storage and database architecture for long-term continuously running spectrum observatories in the US (at Chicago and Blacksburg) and in Turku, Finland. It also describes how the measurement band plan has been collaboratively created taking into account the specifications and limits of the receiving system. Collecting measurements from different geographical locations makes it possible to perform spatial and temporal analysis of the spectrum usage in different parts of the world. As the multi-site measurements are aggregated into one central database, very high requirements are set on the database architecture, equipment, capacity, speed and number of data connections.

Keywords —big data; database; dynamic spectrum sharing; spectrum occupancy measurements; spectrum usage trends

I. INTRODUCTION

The demand for wireless spectrum is constantly increasing as the broadband wireless services are being adopted at an accelerating pace [6]. As the wireless communication technologies themselves are approaching the fundamental theoretical limits of bandwidth efficiency, it is necessary to utilize the existing frequency resources more efficiently. Thus, there is a need for more dynamic access to the spectrum resources, which in turn requires a good understanding of the trends of the current RF spectrum usage in different geographical locations.

The Wireless Networks and Communications (WiNCom) research center at the Illinois Institute of Technology (IIT), with support from the National Science Foundation (NSF), setup the world's first and longest running continuous wideband spectrum observatory [1] based on a dedicated spectrum analyzer in 2007. Over the years, additional FFT-based spectrum sensors were deployed – some of these are wideband, while others monitor select frequency bands. At an international level, IIT in partnership with Virginia Tech, VTT Technical Institute of Finland, Turku University of Applied Sciences, and University of Oulu, has deployed a network of Spectrum Observatories in Chicago and Blacksburg, VA in the USA, and at Turku in Finland. This ongoing Wireless-Finland-US (WiFiUS) Global Spectrum Observatory Network project [16] is funded by the government funding agencies - NSF in the US and Tekes in Finland. This geographically dispersed spectrum observatory

network, with three (3) sensor locations in the US, and two (2) in Finland gathers data of use to policy makers and researchers. The multi-site RF measurement data allows the mapping and modeling of the temporal and spatial variations of the RF environment with sufficiently high fidelity to estimate, characterize, and model spectrum utilization [9]. A central application of the research is to use the knowledge gained to improve the efficiency of spectrum utilization through Dynamic Spectrum Access (DSA) sharing paradigms [7].

RF spectrum measurements from the different sites are currently being collected and aggregated at the central storage database at IIT. In the past, data was stored in discrete files on separate computers. Then a FTP (file transfer protocol) file server was setup for storing all the measurements, but for analysis work the researcher had to download the files to a local workstation. However, since the amount of the data is ever increasing due to addition of new measurement locations and sensors with higher resolution data, this approach is cumbersome while limiting our abilities. In the FTP set up, specific files corresponding to specific dates of interest have to be manually searched and then copied by the user prior to analysis. The ability to query specific dates, times and frequency locations is lacking. Previously a prototype storage architecture [11] with query capabilities was deployed and tested. The system in [11] was implemented on a regular desktop workstation connected to a backend cloud storage hardware system that was limited to 100 Mbps by the local network of the workstation. Since then, IIT has acquired high performance servers and a new storage system with a much higher capacity, located in a server rack adjacent to the IIT spectrum observatories with a 10 Gbps optical fiber connection to the core campus network. Due to the multi-site global dimension of the WiFiUS spectrum observation network, significant design and implementation changes were necessary to the older storage system. The hardware infrastructure additions have also necessitated a complete redesign and reimplementing of the metadata database of global RF measurements.

We introduce a new measurement storage system based on previous work [11] called Tiered Storage for Generic Spectral Data, which coordinates multiple data sources and

types into a unified data format. The system utilizes Cleversafe dsNet object storage [3] as the backend for archival storage; and a metadata/caching database is implemented in PostgreSQL. The PostgreSQL database stores the reference metadata for the measurements to permit queries based on fields such as “dsNet object universally unique identifier” (uuid), date, location and frequency parameters. Cleversafe object storage was chosen due to its long-term reliability with active rebuilding and failure tolerance of drives and nodes within the storage cluster, its ability to expand to multiple global sites, and its capacity scalability to hundreds of petabytes. With the large object store as a backend, measurements, statistics, and analysis results can be archived in the big-data storage system with reduced chance of data loss, and at a lower cost compared to traditional RAID (redundant array of independent disk) or Storage Area Network technologies. Cleversafe [2] estimates that after 100 TB of data reads there is a 99.97% chance of bit errors in a traditional RAID system. With the current dataset of several terabytes, the chances for data loss are significant. With the dsNet object storage system, error checking is done before and after recombination of individual slices to prevent any data loss.

The rest of the paper is organized as follows: Section II describes the WiFiUS measurement setup that has been updated and modified since [16]. The tiered database architecture is described in Section III. Some analysis results using the new measurement data are presented in Section IV. Section V provides benchmark test results for the tiered database, followed by conclusions in Section VI.

II. MEASUREMENT SYSTEM

Spectrum Observatories in both the US and Finland use the CRFS RFeye spectrum measurement receiver [13] as described in [16]. The initial measurement band plan has evolved over time as the project participants have become more familiar with the RF response of the measurement device, and gathered knowledge on what kind of data would be most useful for both regulators and researchers. The new band plan has been collaboratively developed by all of the project participants in different geographical locations and will be used in each of the spectrum observatories to make the direct data comparisons between different sites easier.

The CRFS RFeye has 4 RF external inputs that are connected to an RF switch. The RF switch selects the antenna port to use. The switch then directs incoming signals to the measurement system that has 5 separate RF pre-selectors and front ends, each or covering a different frequency range. The frequency ranges of the frontends are: baseband to 130 MHz, 130-650 MHz, 650-1200 MHz, 1200-4000 MHz and 4000-6000 MHz. The RFeye uses an automatic gain control (AGC) unit with a very fast response time to prevent overload from strong signals coming in through any of the frontends.

The full 30-6000 MHz wideband range is split up into several measurement bands. The obvious starting and

stopping frequencies for a measurement band are the frequency edges of the specific RFeye’s frontend that covers that band. However, the second sweep goes beyond the 650 MHz edge of RFeye’s second frontend to cover the whole TV spectrum of Finland within one band. The band plan information is presented in Table I. The resolution bandwidths have been chosen so as to allow the analysis of the narrowest transmissions of interest within the band.

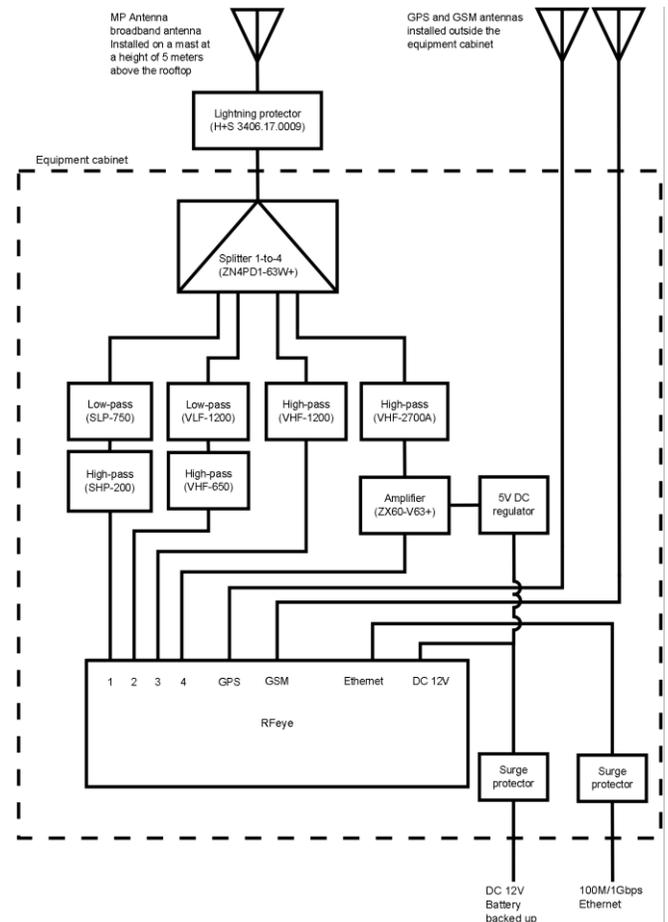


Figure 1. Measurement System Schematic

TABLE I. MEASUREMENT BANDPLAN

Band	Freq. range (MHz)	Resolution bandwidth (kHz)	Scan interval (seconds)	RF Input Port
1	30-130	78.125	10	1
2	130-800	39.0625	3	1
3	650-1200	39.0625	3	2
4	1200-3000	39.0625	3	3
5	3000-6000	78.125	3	4

Figure 1 describes the RFeye measurement system - the device itself, the connected filters, amplifier and antennas. A

multi-polarized broadband antenna is used to measure the RF spectrum. The antenna signal is split into the four RF inputs of the RFeye spectrum measurement receiver. Band-pass filters are applied prior to the RF inputs to attenuate the strong out-of-band transmissions. Otherwise the strong signals cause the RF front end to use a high attenuation setting, which raises the overall noise floor while lowering the sensitivity to weak signals. This was a problem in the previous setup [16], where only one RF port was used and there was no bandpass or band-reject filtering. The use of different filters feeding into the 4 different ports of the RFeye eliminates this problem.

Strong FM radio signals are attenuated with a 200 MHz high-pass filter in RF input 1, but the measured FM powers are still strong enough to permit easy detection. A 750 MHz low-pass filter is installed before RF input 1 to attenuate the strong GSM signals above 750 MHz. Since there is a RF front end edge at 130 MHz, the RF input 1 signal is employed for two different sweeps - Bands 1 and 2 (see Table I). As mentioned above, the second sweep goes beyond the 650 MHz edge of RFeye's second frontend to cover the whole TV spectrum of Finland within one band. This facilitates the analysis of TV White Space devices and PMSE (Program Making and Special Events) operating in the 470-790 MHz region. A sweep from one RF eye frontend edge to the next is performed in the RF input 2, and the band-pass filters are chosen accordingly to attenuate the out-of-band transmissions.

The strong GSM (Global System for Mobile Communications) signals below 1200 MHz are attenuated with a high-pass filter in RF input 3. All of the stronger signals at lower frequencies are attenuated with a 2700 MHz high-pass filter in RF input 4. As there is a multitude of interesting but weak signals on the 3 to 6 GHz band, a low-noise amplifier (LNA) is installed to strengthen these signals. Using the 1-to-4 splitter adds 6-7 dB loss, which makes an amplifier a necessity for the signals beyond 3 GHz. The typical gain of the LNA is 15.4 to 19.2 dB within the measured band. Sweeps for Band #5 goes beyond the 4 GHz edge of the fourth frontend (1200-4000 MHz), but this introduces only a minor change in the noise floor, which can barely be seen in Figure 4. The band-pass filters introduce relatively low insertion losses, usually below 1 dB. Data sheets with detailed information about the band-pass filters, splitter and the amplifier can be found from the manufacturer web-site [8] by using the part numbers listed in Figure 1.

III. REDESIGNED 3-TIER DATA STORAGE SYSTEM

As measurement efforts have expanded, the ability to manipulate the dataset has grown increasingly complex, especially in regards to long-term time series analysis. Measurements are stored in different data formats, depending on the sensor used, and analysis requires manual intervention by the researcher to find the relevant files and import them to their local machine. In order to better

manipulate and store long-term spectrum measurements, a new storage methodology had to be designed which allows for larger storage capacities and ensures reliability of the data. We introduce a storage methodology called Tiered Storage of Generic Spectral Data (TSGSD), which uses a traditional relational database for measurement metadata, Cleversafe dsNet Simple Object Storage for measurement storage, and a caching layer for optimal retrieval speeds.

With internationally distributed collection sites, analysis requires centralization of data. See Appendix Figure A1 for the data synchronization path. Using the TSGSD at IIT, researchers are able to coordinate measurement data from distributed sites. Through the use of identical sensor configurations (hardware and software measurement configurations), direct comparisons between spectral data are feasible.

TSGSD consists of a computing cluster containing the dsNet storage system (76 TB of raw storage), dedicated Accesser devices for object retrieval from the dsNet storage system, the database server for metadata, an analysis server, and frontend web servers for web retrieval of data and plots. The cluster is located on the same server-rack to maximize data retrieval speeds. The whole setup is then connected via fiber-optic trunk lines to IIT research labs and to the campus Internet connection.

One of the central features of the TSGSD is the object storage daemon (OSD). The OSD is a background process on the metadata server, which coordinates the preprocessing of incoming raw measurement files from multiple sources to output two data streams – one for the metadata and the other for measurements.

Aggregation operations are performed upon request by a frontend server, and then stored in the TSGSD. By performing aggregation operations within the cluster ensures that later operations on the same aggregated dataset are stored within the dsNet as a reduced set of measurements, instead of forcing retrieval of large quantities of un-aggregated data. Typical aggregation operations are hourly, daily, and weekly max-hold or average within a band or across the entire spectrum. Aggregation operations occur at the resolution of the measurement, to ensure visibility of narrow-band signals. The aggregation type depends on the analysis to be performed, and can be input by the researcher. Appendix Figure A2 describes the path for a typical retrieval operation of data from TSGSD.

Central to the TSGSD is a metadata database, which correlates object identifiers (unique uuid identifiers of measurement data-blocks in the dsNet system) and relevant measurement metadata, including site information (GPS coordinates, fixed or mobile measurements), date and times, scan ranges, and aggregation information.

The metadata structure is designed keeping in mind the need to support both raw data scans, and data analysis results. Once a data set is aggregated after analysis, the

reduced dataset is inserted into TSGSD for future use. Retaining copies of pre-analyzed data minimizes the need for repeated retrievals of raw data. Typical dataset for analysis consists of 3.2 GB of raw data per day per site. Queries for aggregated data are performed against the metadata database, and partial or full matches to a query return a previously processed copy. In order to further improve the rate of data analysis, a subset of the objects stored in the dsNet is cached in an in-memory binary table component of the database, for immediate retrieval.

Integration with existing raw data in various formats is an important function, and the primary motivation for using a large archival object store as a backend is the ability to scale storage capacity as the data capture rate increases. By preprocessing multiple data sources into a single generic format, analysis code can be simplified and reused. The original source of the data is maintained in the metadata database, which provides important information about the specifications of the sensors and equipment used to measure the spectrum.

Existing research by IIT WiNCom [11] using dsNet object storage had a few limitations: a 100 Mbps LAN link bottleneck, limited archive size on shared storage (4 TB). [11] also made use of a different encoding scheme which current research has shown to be more expensive in terms of storage utilization and processing time. The new TSGSD system has direct access to campus fiber trunk lines, with direct connections to analysis and database servers, making queries into the dsNet faster. The metadata database also uses traditional relational database design, in comparison to the MongoDB [10] document store used in [11]. This allows for more information about individual objects to be stored with less redundancy. In [11], using MongoDB and Binary Javascript Object Notation (BSON) encoding allowed scan data to be exported and stored directly as an object.

However, when using SQL (structured query language) variants such as in TSGSD, the overhead for BSON encoding is less advantageous. By using the LZ77 compression algorithm (used in gzip [4]), scan-data encoding time is reduced by a factor of 10, and the stored object size reduced by 20% compared to BSON.

IV. ANALYSIS RESULTS

This section presents occupancy measurement results from Turku, Finland, as well as some of the web browser based analysis and live monitoring tools. The measurements are conducted with (a) the band plan introduced in [16] and (b) the band plan introduced in Table I of this paper. The occupancy results with the old band plan were obtained from the 100-6000 MHz band measurement in [16] with a frequency resolution of 312.5 kHz, while the new band plan results are obtained with the parameters listed in Table I.

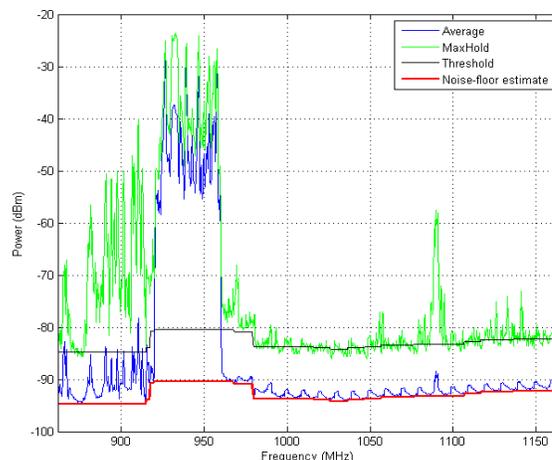


Figure 2. Calculated noise floor and chosen threshold in Turku, 22nd of November 2013.

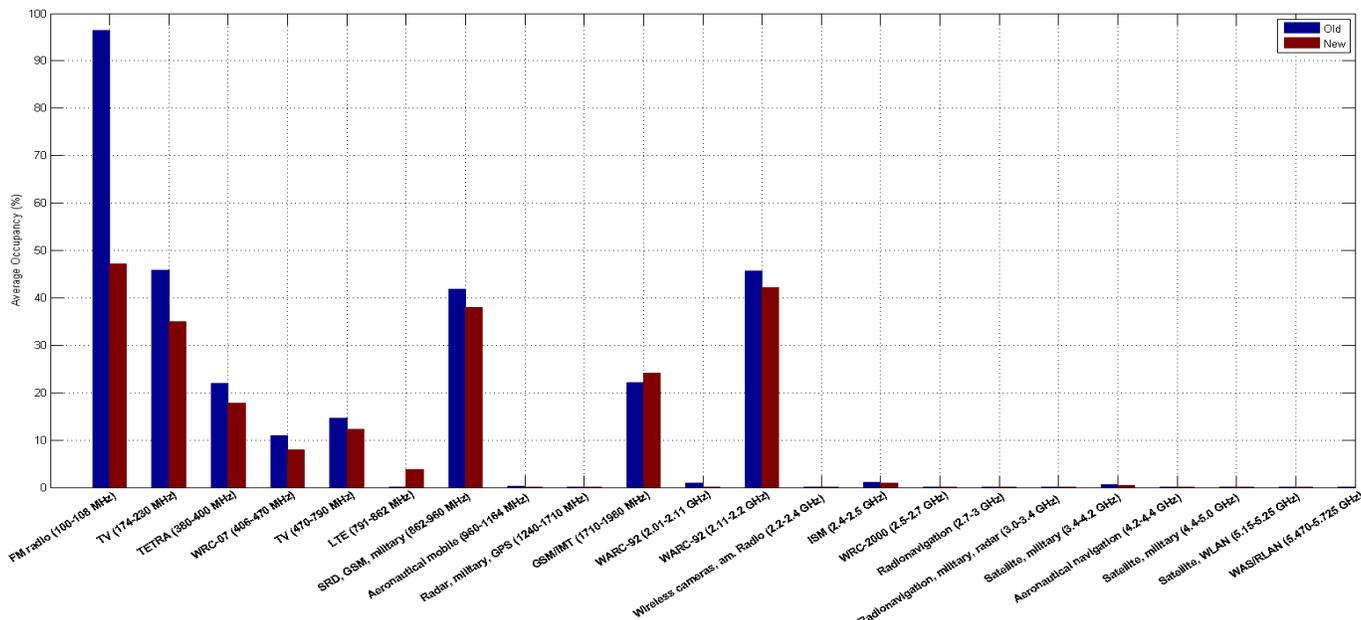


Figure 3. Occupancy estimates across different bands using the older 2013 bandplan (blue) and newer 2014 bandplan (red)

To determine occupancy, we need to calculate the noise floor for each of the frequency points and add a threshold so as to determine if the measured power exceeds a minimal power level and thus is a valid signal [14]. An algorithm based on [12] was developed at IIT to automatically estimate the noise floor. A more detailed description of how the algorithm functions is found in [16]. The noise floors were automatically calculated, but the threshold offsets for the occupancy band chart were manually chosen by carefully examining each band. Figure 2 shows the resulting noise floor estimate from the algorithm, maximum and average power spectrum, and the corresponding threshold used to determine the occupancy. The threshold for this specific band was chosen to be 10 dB above the noise floor estimate.

The average occupancy statistics are presented in Figure 3. Approximately one day’s worth of data was used to generate the bars in the chart. The old band plan data is from November 22, 2013 and the new band plan data is from March 5, 2014. The labels at the bottom of the chart indicate some of the wireless services in that band in Finland and the start and stop frequencies. As the data from the new band plan has higher resolution, the signals can be better distinguished and more accurate results can be achieved. An example of this is seen from the FM radio band occupancy, where the old measurement system data inaccurately gave a high number. This was because the 312.5 KHz resolution was too low to distinguish the adjacent narrow 200 KHz FM channels from each other, and the unused FM channels could not be resolved. Another band where we see a change in the occupancy is in the 800 MHz LTE band. This is because in Finland, 800 MHz band LTE transmissions began only in the beginning of year 2014. In the higher frequency bands, as shown in Figure 4 for the 3-6 GHz region, we do see signals in the plot of max-hold power spectrum, but they are not discernable in the average power spectrum plot. This means that the signals in the 3-6 GHz regions though detectable have low duty cycles. This explains why the occupancy bar chart of Figure 3 shows very low occupancy in those bands. Very low duty cycle naturally results in a very low occupancy when averaged over the 24-hour period used to calculate occupancy.

Figure 4 shows a comparison of max-hold power spectrum in the higher frequency band from 3 to 6 GHz with the old and new band plans. The careful selection of the edges for Band #5, the amplifier and the high-pass filter have flattened the fluctuations (due to RFeye AGC activity) in the noise floor of the new system. This has also improved the sensitivity of the system considerably. For example, military radar activity which was not visible before can clearly be seen on the 3000 to 3200 MHz band. One of the benefits of the spatial measurements is that we can compare RF spectrum usage across different geographical locations. In the Chicago measurement locations, there is almost no detected activity in the 3000-3200 MHz range.

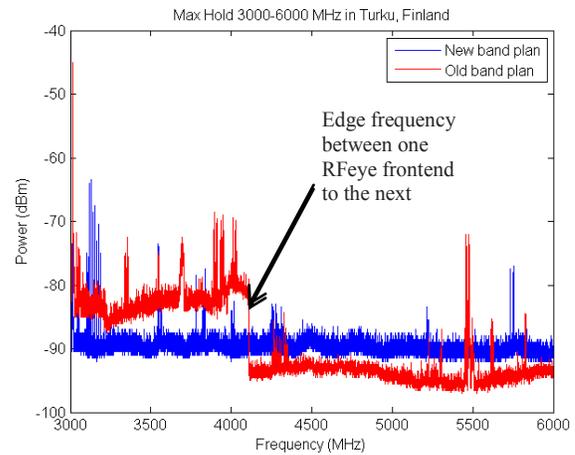


Figure 4. Maximum power spectrum in Turku for 3 to 6 GHz band for the analyzed days.

The results presented so far were obtained locally at a workstation running MATLAB. Remotely, the spectrum observatory data can also be queried via a web browser based user interface, where the user does not need to download the data or install any software for analysis other than a modern web browser. The user can do different types of analyses, such as power spectrum plots, spectrograms and time series for the time span and frequencies chosen. The plots can be interactively zoomed and clicked on to provide numerical data for a specific point in the plot. Currently this tool cannot directly use the data from the RFeye measurements, but that capability is being worked on. Figure 5 shows an example plot of the browser based analysis tool. The input data is from a different spectrum observatory system [1] where HDF5 files [5] are in use. All of the spectrum observatories also provide a live web-stream which permits live spectrum monitoring. The streams are aggregated into a server-based web-page that displays the current RF measurements being made at all locations.

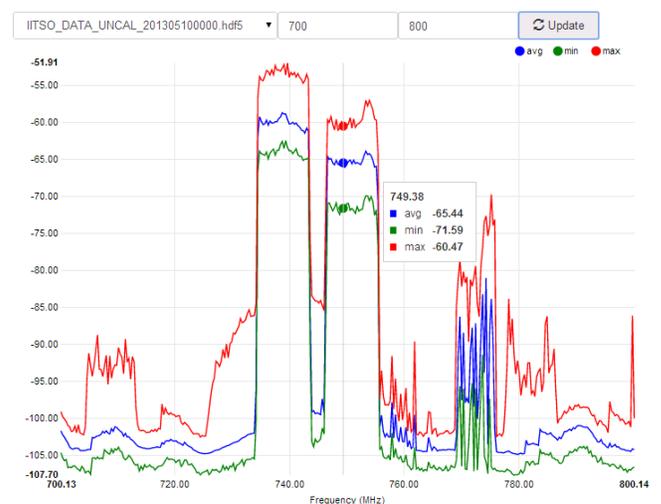


Figure 5. Web-Browser based Analysis and Plotting of Spectrum data

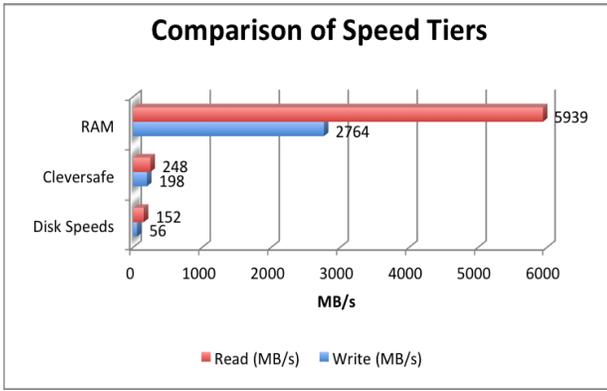


Figure 6. Comparison of Speed of Individual Storage Tiers

V. PERFORMANCE OF TIERED STORAGE WITH GENERIC SPECTRAL DATA

In designing a new data storage system for long-term use, three major factors were of interest: cost, scalability, and reliability. Using a single RAID server with a large number of drives would provide similar total capacity, but with significant storage overhead, a single point of failure and the speed limitations of individual drives. Traditional SANs were deemed to be too expensive and difficult to scale to big-data capacity. Existing research [11] at WinCom using object storage provided insight into the scalability, redundancy, and cost effectiveness of object based storage. Increasing the ability to perform greater analyses on spectral data requires greater retrieval speeds. Figure 6 shows a comparison of the different tiers used in TSGSD. Figure 6 shows the maximum speed of each tier – the RAM for the cache table, the disk for the metadata database, and dsNet for the object store.

By taking advantage of available memory in the database server, a cache of approximately 64GB of recent objects can be maintained (Approximately 3 days’ worth of full-resolution measurements). This allows direct analysis without the need for disk access.

Another feature of TSGSD is the reduction of distance between data storage and analysis servers, which eliminates data transfer times. Custom software developed at WinCom for preprocessing binary RFeye files executes faster than vendor-provided software, and physical locality of data eliminates file transfer times. The elimination of data transfer times reduces the overall processing time by 18%. Using the multi-core processor of the database server, up to 6 pre-processing threads can be spawned concurrently via multi-processing techniques to cut down data insertion times 6 fold. The object-storage-daemon runs in the background on the database server, and processes synchronized measurements on arrival; eliminating the need for researchers to wait for file preprocessing. Table II compares the wait times for pre-processing data using a standalone workstation versus the TSGSD system.

TABLE II. DATA PROCESSING TIMES COMPARED

Processing Times for single day data	Preprocessing Manufacturer Software at local workstation	TSGSD with Custom Software
Transfer Wait Time	307 s	0.00 s
Preprocessing Time	1645 s	1638.4 s

The metadata database operations are expected to be mostly selection operations, with insertion operations reserved for raw measurement data and for storing pre-analyzed results. For queries and/or insertions from multiple sources, the database server supports hundreds of connections. At 100 concurrent connections, the database server can support 67327 selection operations per second. With a query involving 8 search fields, this provides access to 8415 individual objects per second. Assuming each query represents an individual uncached 4 MB object, the queries would saturate the dsNet object storage link speed of 2x1 Gbps. This means the only bottleneck is the link speed between the database server and dsNet. It is technically easy to expand the network link capacity by acquiring additional dsNet accessor devices. Crucially, neither the metadata database nor the dsNet disk speed is a bottleneck for the TSGSD system in this design.

By using LZ77 compression through the gzip python library, JSON data is stored in a compressed format in TSGSD and can be read by analysis software sweep-by-sweep, without requiring conversion of the entire file. The compression reduces intermediate file sizes by 79%, and allows a greater amount of measurements and statistics to be stored at the caching layer. The various storage format sizes for a typical days’ worth of spectral measurements from a single site are shown in Table III.

TABLE III. PROCESSED DATA SIZES

Processed and Raw Dataset Sizes	Local workstation with vendor-provided pre-processor	TSGSD using gzip
Data Size:	3276.8 (MB)	3276.8 (MB)
Processed size:	24985.6 (MB)	5177.3 (MB)

VI. CONCLUSION

The new band plan for the spectrum measurement system allows collection of spectrum data with higher sensitivity through careful use of band-pass filters to attenuate strong out-of-band transmissions, and by using an amplifier for the weaker transmissions in the 3-6 GHz band. The earlier band plan did not take into account the RF response of the measurement device, which thus caused the noise floor to fluctuate within the measurement sweeps. This problem has been eliminated by the new band plan.

The expansion of spectrum measurement data collection in multiple sites at higher resolutions has required the use of more complex computing infrastructure to store and analyze vast amounts of data efficiently. Traditional methods for storage and analysis have reached their effective limits, and TSGSD presents a method of archival storage combined with fast query capabilities and a high-speed network.

Through the development of software on top of the system presented in TSGSD, the results of analysis operations that occur at multiple sites will be stored to save time in further research involving more complex analysis tools. The database components within TSGSD scale through replication, and the dsNet can later be expanded to multiple geographically distributed sites. Through careful system design, dsNet objects can be stored redundantly at multiple sites, where read operations would only activate closest local-site nodes.

The on-server analysis and aggregation system in TSGSD allows for high-performance international research, by centralizing the analysis work near the data storage system. Centralized aggregation and analysis near the data system ensures only the minimum required dataset is sent to international researchers over the Internet.

During 2014 the spectrum occupancy data collection will be enhanced even further. Portable “backpack” measurement unit will be utilized to provide complementary information on the utilization of specific systems, and the most interesting frequency bands from the dynamic spectrum access point-of-view. It is still under debate whether a few fixed measurement locations are sufficient to produce spectrum occupancy evaluations that could be used a basis for radio frequency regulation. The portable unit can be used to focus on detailed spatial variations. Currently, one of the most interesting frequency bands in the Europe is 2,3-2,4 GHz, which is the first the first band considered for licensed shared access (LSA). The portable measurement system will provide insight for example how wireless cameras are used in the field.

Through continued development of in-browser plotting combined with the on-site aggregation of data, researchers at IIT, VT, VTT, Oulu University and Turku University of Applied Sciences can quickly analyze and view features of the data collected from all the measurement sites. Moving forward, the computer infrastructure presented in this paper will be a central resource and key enabler of future international research undertaken in this WiFiUS project.

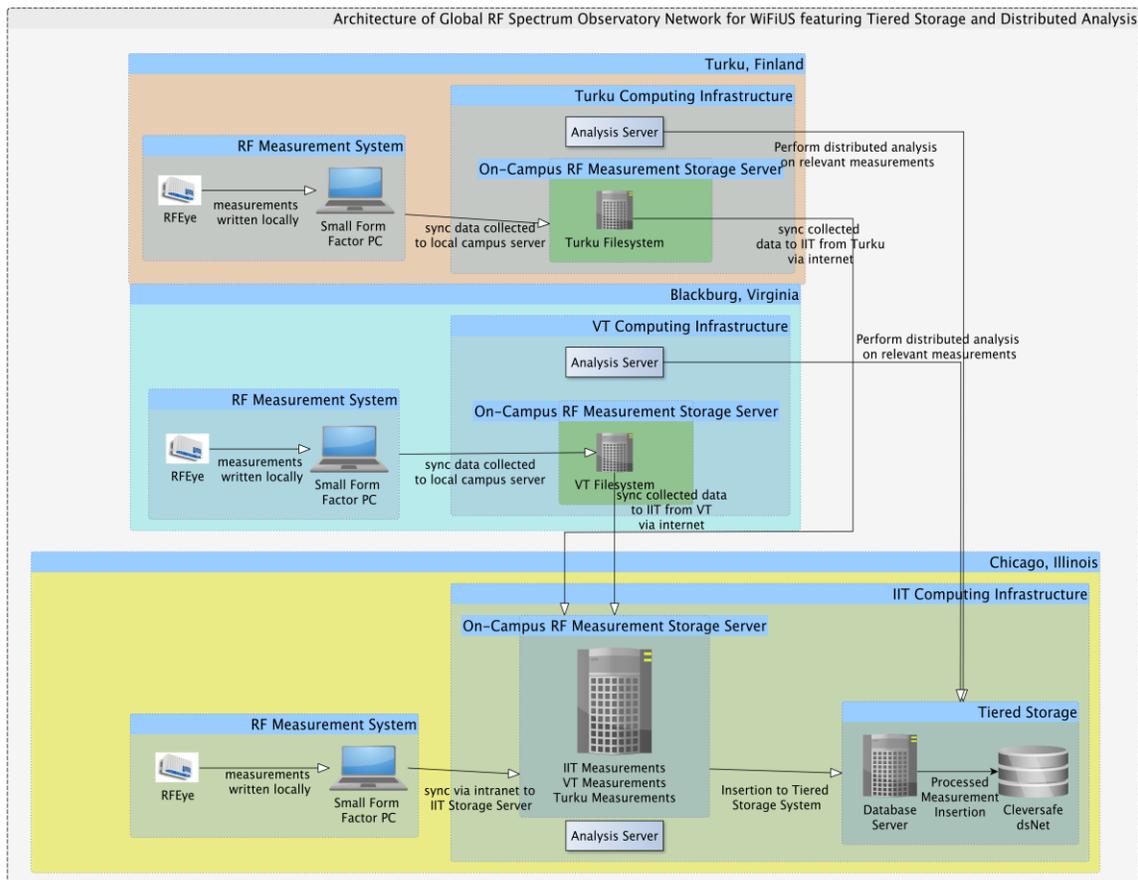


Figure A1. Collection methodology for spectrum measurements

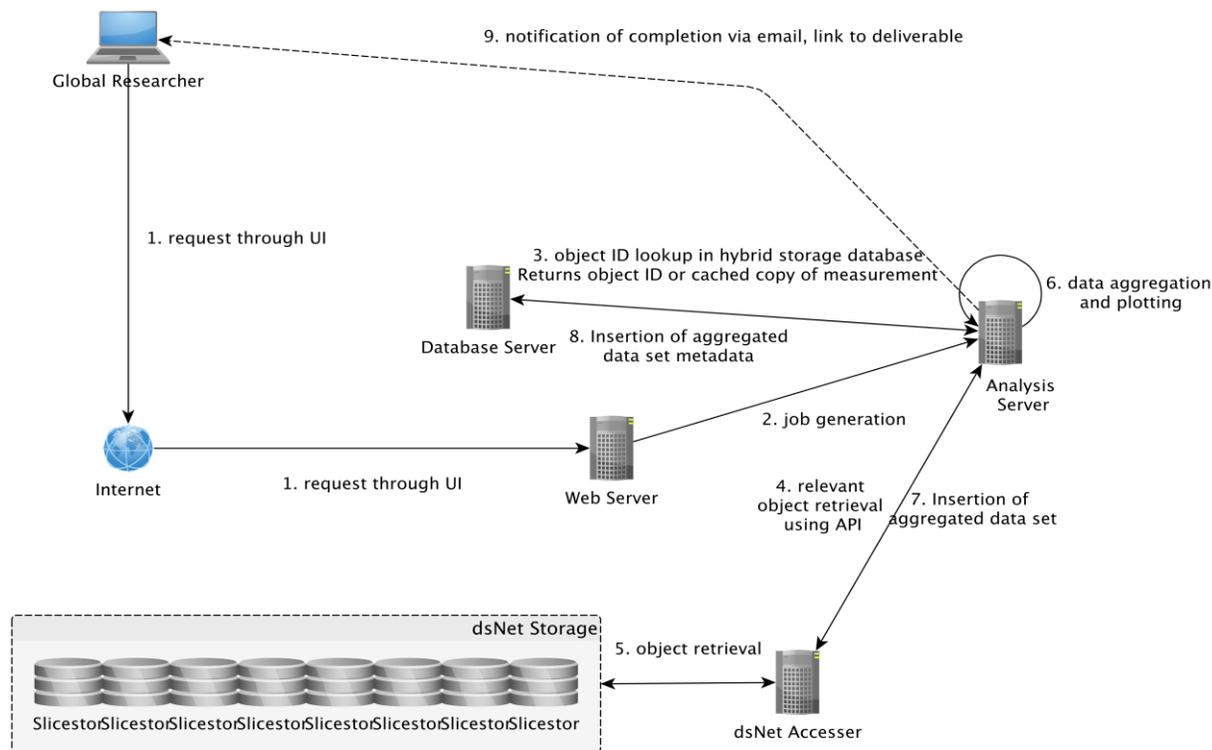


Figure A2. Retrieval operations for spectrum measurement

APPENDIX

See Figure A1 for an overview reference of the data collection system. See Figure A2 for block diagrams of retrieval operations in the TSGSD.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the National Science Foundation (NSF) and Tekes. This research is supported by the NSF under Grant NSF EAGER CNS 1265351. The project partners in WiFiUS Global Spectrum Observatory Network project have contributed significantly to the development of the new system with the updated band plan, especially at Turku University of Applied Sciences. The authors are grateful to Cleversafe for the kind donation of a dsNet Storage system and their support in our research efforts.

REFERENCES

- [1] Bacchus, R.B.; Fertner, A.J.; Hood, C.S.; Roberson, D.A., "Long-Term, Wide-Band Spectral Monitoring in Support of Dynamic Spectrum Access Networks at the IIT Spectrum Observatory," *3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2008*. Chicago, IL, Oct. 08
- [2] Cleversafe Inc. *Why RAID is Dead for Big Data Storage*. Web. <<http://www.cleversafe.com/library/whitepapers/register-to-download-raid-is-dead>> (Registration required.)
- [3] Cleversafe Inc. Web-site: Retrieved March 2014 <http://cleversafe.com>
- [4] GNU Gzip. *Overview*. Web. Retrieved May 25 2014. <http://www.gnu.org/software/gzip/manual/html_node/Overview.html#Overview>
- [5] *The HDF Group – Information, Support and Software*. Web. Retrieved 24 May 2012. <<http://hdfgroup.org>>
- [6] Report ITU-R M.2243. Assessment of the global mobile broadband deployments and forecasts for International Mobile Telecommunications. 2011.
- [7] V. Kone, L. Yang, X. Yang, B. Y. Zhao and H. Zheng, "The effectiveness of opportunistic spectrum access: A measurement study," *IEEE/ACM Transactions on Networking*, vol. 20, pp. 2005-2016, December 2012.
- [8] Mini-Circuits Web-site: Retrieved March 2014 <<http://www.minicircuits.com>>
- [9] M. McHenry, D. McCloskey, G. Minden and D. Roberson, "Multi-band multi-location spectrum occupancy measurements," in *Proc. ISART*, March 2006.
- [10] "MongoDB." MongoDB. 10 Gen, Inc. Web. Retrieved March 2014 <<http://mongodb.org>>
- [11] Noorts, G; Engel, J; Taylor, J; Bacchus, R; Taher, T; Roberson, D; Zdunek, K;, "An RF Spectrum Observatory Database based on a Hybrid Storage System," *IEEE Dynamic Spectrum Access (Dyspan) conference at Seattle, WA, October 2012*.
- [12] Ready, M., Downey, M. and Corbalis, L. 1997. Automatic noise floor spectrum estimation in the presence of signals. (877-881).
- [13] CRFS RFeye. Web-site, accessed March 2014: <<http://www.crf.com/products/rf-sensor-rfeye-node/>>
- [14] Report ITU-R SM.2256. Spectrum occupancy measurements and evaluation. 2012.
- [15] A. Silberschatz, *Database System Concepts*. Sixth Edition. New York, NY. McGraw-Hill 2011.
- [16] Taher, T; et al., "Global spectrum observatory network setup and initial findings," to be presented at ICST Cognitive Radio Oriented Wireless Networks Conference (CROWNCOM) at Oulu, Finland, June 2014.