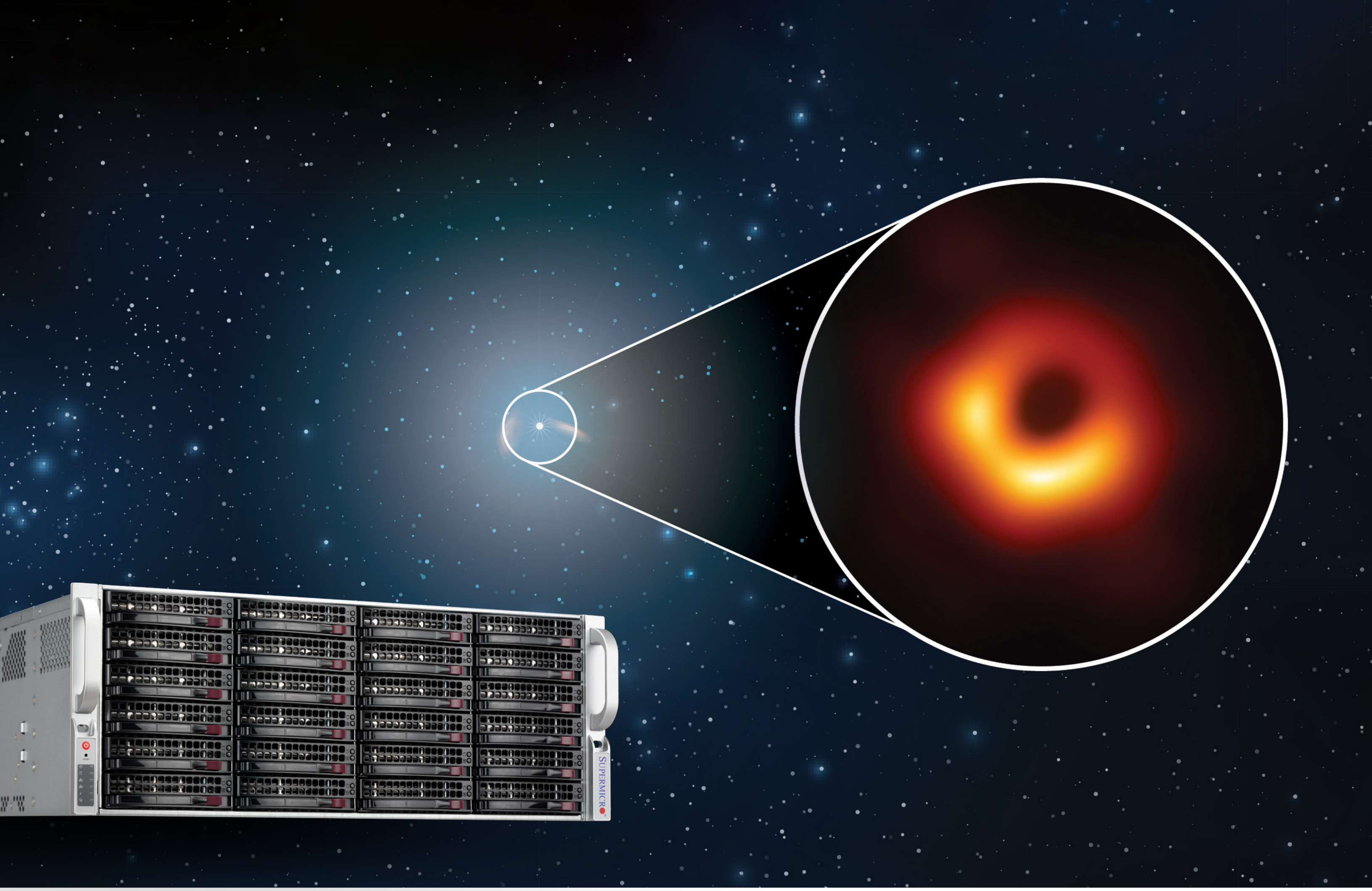


Table of Contents

White Paper

Data Processing and Storage for Black Hole Event Horizon Imaging

Tom Coughlin, President, Coughlin Associates, Inc.

Executive Summary

The first imaging of the event horizon for a black hole involved an international partnership of 8 radio telescopes with major data processing at MIT and the Max Planck Institute (MPI) in Germany. The contribution of the brilliant scientists was aided by the expanded computing power of today's IT infrastructure. Processing of the 4 petabytes (PB) of data generated in the project in 2017 for the original imaging utilized servers and storage systems, with many of these servers coming from Supermicro. Figure 1 shows the MIT Correlator Cluster as it looked in 2016.

Super Micro Computer, Inc.
980 Rock Avenue
San Jose, CA 95131 USA
www.supermicro.com

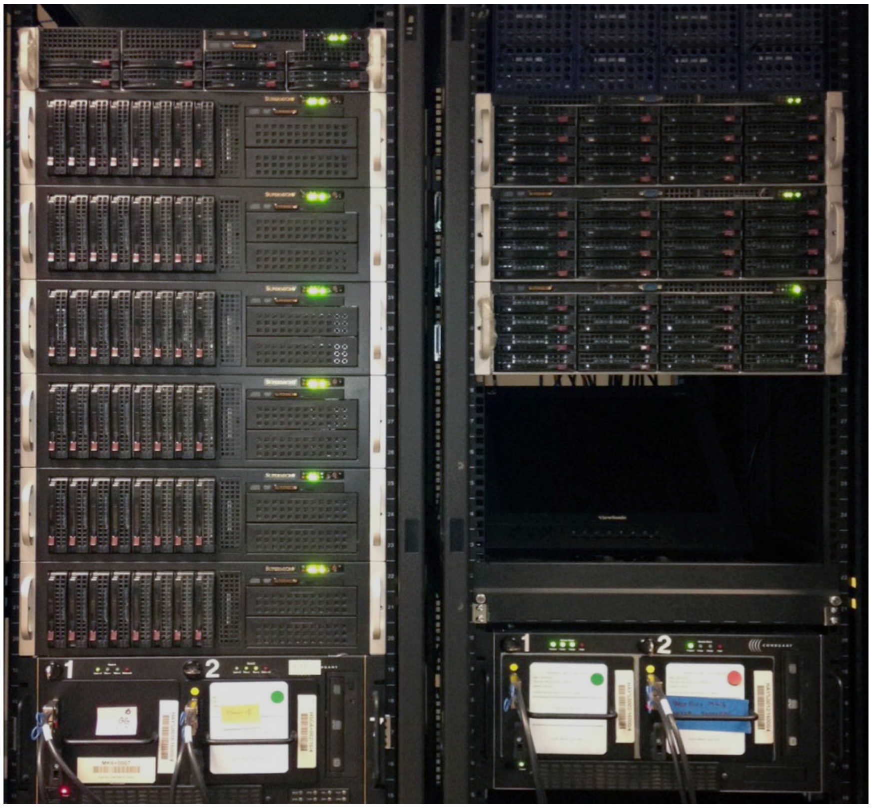


Figure 1. *Detail of DiFX Correlator cluster racks showing Supermicro servers*[1]

Black Holes and their Event Horizons

With the publication of Albert Einstein's general theory of relativity in 1915, our views of the nature of space, time and gravity were changed forever. Einstein's theory showed that gravity is created by the curvature of space around massive objects.

When the nuclear fusion processes that create the radiation of stars begins to run out of fuel and no longer exert sufficient outward pressure to overcome the gravity of the star, the star's core collapses. For a low mass star like our Sun, this collapse results in a white dwarf star that eventually cools and becomes a nearly invisible black dwarf. Neutron stars are the result of the gravitational collapse of the cores of larger stars, which just before their collapse, blow much of their mass away in a type of supernova. Even larger stars, more than double the mass of our sun, are so massive that when their nuclear fuel is expended, they continue to collapse, with no other force able to resist their gravity, until they form a singularity (or point-like region) in spacetime. For these objects, the escape velocity exceeds the speed of light and hence conventional radiation cannot occur and a black hole is formed.

[1] Titus et. al., Haystack Observatory VLBI Correlator 2015–2016 Biennial Report, in International VLBI Service for Geodesy and Astrometry 2015+2016 Biennial Report, edited by K. D. Baver, D. Behrend, and K. L. Armstrong, NASA/TP-2017-219021, 2017

The black hole singularity is surrounded by a region of space that is called the event horizon. The event horizon has a strong but finite curvature. Radiation can be emitted from the event horizon by material falling into the black hole or by quantum mechanical processes that allow a particle/antiparticle pair to effectively tunnel out of the black role releasing radiation (Hawking radiation).

Black holes are believed to be fairly common in our universe and can occur in many sizes, depending upon their initial mass. Hawking radiation eventually results in the "evaporation" of a black hole, and smaller, less massive, black holes evaporate faster than more massive black holes. Super massive, long-lived black holes are believed to be at the center of most galaxies, including our Milky Way. Until 2019, no one had imaged the space around a black hole.

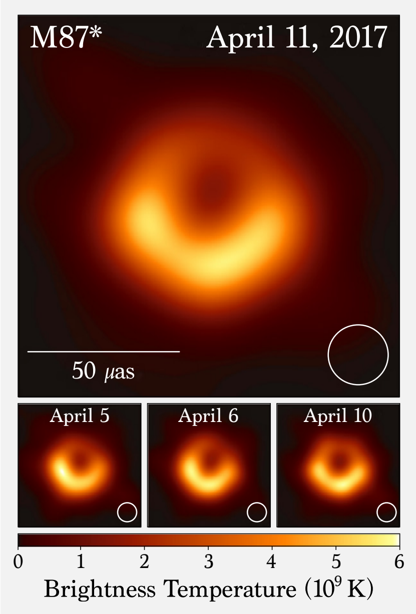
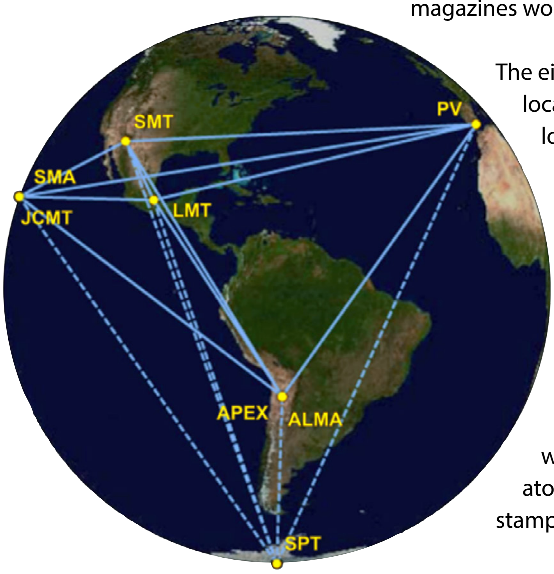


Figure 2. *First M87 Event Horizon Telescope Results. III. Data Processing and Calibration, The Event Horizon Telescope Collaboration, The Astrophysical Journal Letters, 875:L3 (32pp), 2019, April 10*

The Black Hole Event Horizon Imaging Project

Radio telescopes allow observations even in the presence of a cloud cover and microwave radiation is not absorbed by interstellar clouds of dust. For these reasons, radio telescopes provide more reliable imaging of many celestial objects. 1.3mm wavelength is a popular observing frequency.

Using an array of 8 international radio telescopes in 2017, astrophysicists used sophisticated signal processing algorithms and global very long baseline interferometry (VLBI) to turn 4 petabytes of data obtained from observations of a black hole in a neighboring galaxy into the first image of a black hole event horizon. This particular black hole is located 55 million light-years away (in galaxy Messier 87, M87). It is 3 million times the size of the Earth, with a mass 6.5 billion times that of the Earth's sun. A version of the images shown in Figure 2, showing a glowing orange ring created by gases and dust falling into the black hole, appeared on the front pages of newspapers and magazines worldwide in mid-April 2019.



The eight different observatories, located in six distinct geographical locations, shown in Figure 3, formed the Event Horizon Telescope (EHT) array. This collection of telescopes provided an effective imaging aperture close to the diameter of the earth, allowing the resolution of very small objects in the sky. Observations were made simultaneously at 1.3mm wavelength with hydrogen maser atomic clocks used to precisely time stamp the raw image data.

Figure 3. *Akiyama et. al., First M87 Event Horizon Telescope Results. III. Data Processing and Calibration, The Event Horizon Telescope Collaboration, The Astrophysical Journal Letters, 875:L3 (32pp), 2019, April 10*

Capturing and Storing EHT Data

VLBI allowed the EHT to achieve an angular resolution of 20 micro-arcseconds, said to be good enough to locate an orange on the surface of the Moon, from the Earth. Observation data was collected over five nights, from April 5–11, 2017. These observations were made at each site as the weather conditions were favorable. Each telescope generated about 350 TB of data a day and the EHT sites recorded their data at 64 Gb/s.

The data was recorded in parallel by four digital backend (DBE) systems on 32 helium-filled hard disk drives (HDDs), or 128 HDDs per telescope. So, for the 8 telescopes, 1,024 HDDs were used. The Western Digital helium filled HDDs used were first obtained in 2015, when these were the only He-filled sealed HDDs available. Sealed He-filled HDDs were found to operate most reliably at the high altitudes of the radio telescopes. Processing of the collected data was simultaneously done using correlators at the MIT Haystack Observatory (MIT) and the Max Planck Institute in Germany (MPI).

According to Helge Rottmann from MPI[2] "For the 2017 observations the total data volume collected was about 4 PB." He also said that starting in 2018 the EHT collected data doubled to 8 PB. The DBEs acquired data from the upstream detection equipment using two 10 Gb/s Ethernet network interface cards at 128 Gb/s. Data was written using a time sliced round-robin algorithm across the 32 HDDs. The drives were mounted in groups of eight in four removable modules. After the data was collected the HDD modules were flown to the Max Planck Institute (MPI) for Radio Astronomy in Bonn, Germany for high frequency band data analysis and to the MIT Haystack Observatory in Westford, Massachusetts for low frequency band data analysis.

Vincent Fish from the MIT Haystack Observatory said that[3], "It has traditionally been too expensive to keep the raw data, so the disks get erased and sent out again for recording. This could change as disk prices continue to come down. We still have the 2017 data on disk in case we find a compelling reason to re-correlate it, but in general, once you've correlated the data correctly, there isn't much need to keep petabytes of raw data around anymore."

The EHT Correlators

The real key to extracting the amazing images of the event horizon of a black hole was the use of advanced signal processing algorithms to process the data. Through the receiver and backend electronics at each telescope, the sky signal is mixed to the baseband, digitized, and recorded directly to hard disk, resulting in petabytes of raw VLBI voltage signal data. The correlator uses an a priori Earth geometry and a clock/delay model to align the signals from each telescope to a common time reference. Also, the sensitivity of the antennas had to be calculated to create a correlation coefficient between the different antennas.

---

[2] Email from Helge Rottmann, Max Planck Institute, May 7, 2019
[3] Email from Vincent Fish, MIT Haystack Observatory, April 30, 2019

Figure 4. *Photo by Nancy Wolfe Kotary, MIT Haystack Observatory*

The actual processing was performed with DiFX software[4] running on high performance computing clusters at MPI and MIT. The clusters are composed of 100s of servers, thousands of cores, high performance networking (25 GbE and FDR Infiniband) and RAID storage servers. The MIT Haystack correlator is shown in Figure 5.

The MPI Cluster, located in Bonn Germany, is comprised of 3 Supermicro head node servers, 68 Compute nodes (20 cores each = 1,360 cores), 11 Supermicro storage RAID servers running BeeGFS parallel file system with a capacity of 1.6 petabytes, FDR Infiniband networking, 15 Mark 5 playback units, as shown in figure 5, and 9 Mark 6 play back units.

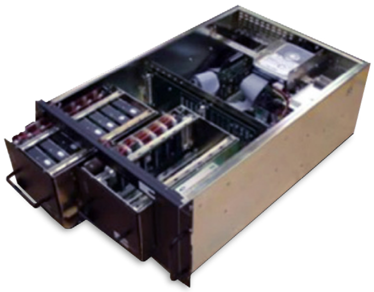


Figure 5. *Mark 5 Playback Unit*

The MIT Cluster is housed within 10 racks (9 visible in Figure 4). Three generations of Supermicro servers were used with the newest having two 10-core Intel® Xeon® CPUs. The network consists of Mellanox® 100/50/40/25 GbE switches with the majority of nodes on the high speed network at 25GbE or higher Mellanox PCIe add-on NICs. In addition to the Mark 6 recorders there is half a petabyte of storage scattered across the various Supermicro storage servers for staging raw data and archiving correlated data product1.

Supermicro in the EHT Correlators

The Correlators make extensive use of Supermicro's large portfolio of Intel Xeon Processor based systems and Building Block Solutions® to deploy fully optimized compute and storage solutions for various workloads. For example, the Haystack DiFX Correlator depicted in Figure 6 leverages Supermicro solutions for compute, storage, administration, and maintenance tasks.

---

[4] Deller, A. T., Brisken, W. F., Phillips, C. J., et al. 2011, PASP, 123, 275

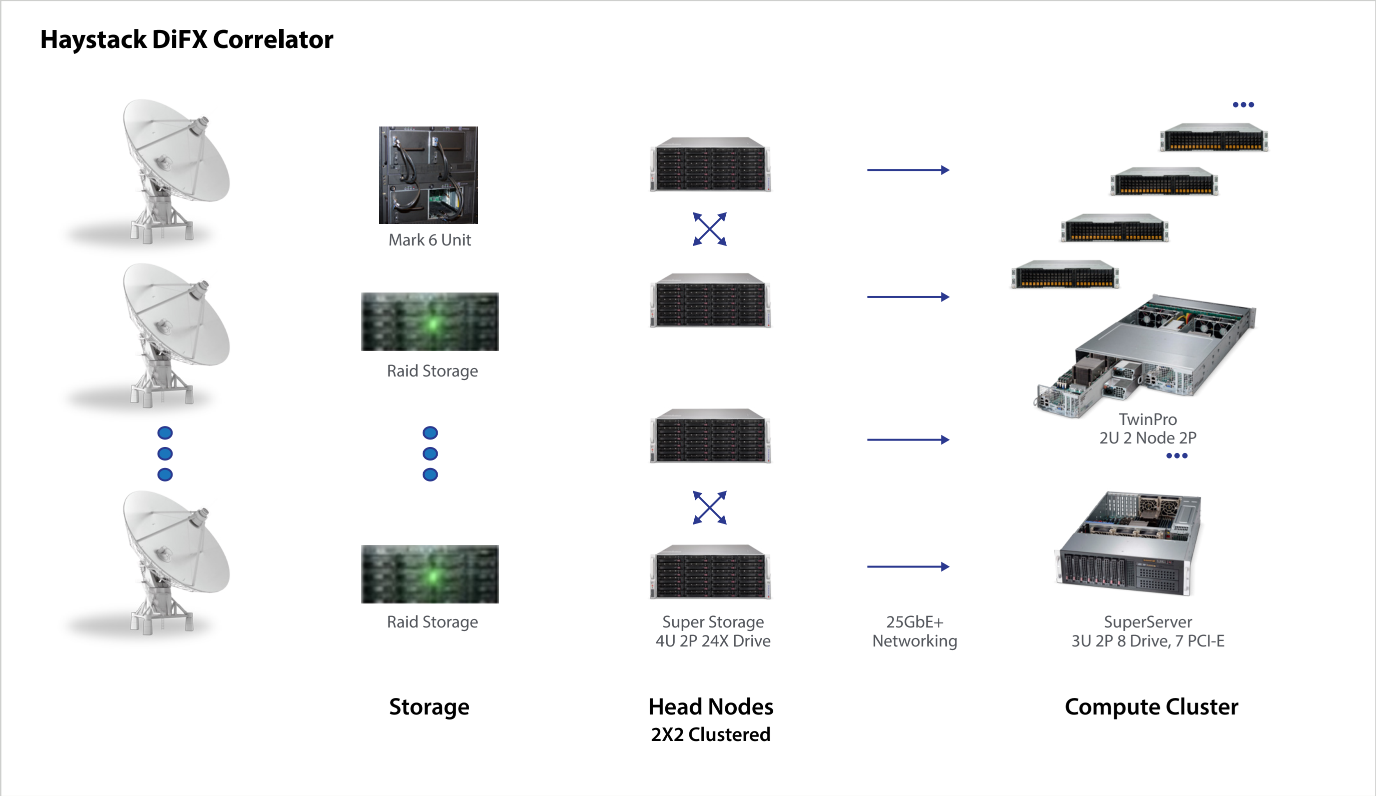


Figure 6. *Haystack DiFX Correlator*

Due to the high resource demand of DiFX, 2x2 clustered Supermicro headend nodes (4 total), shown in Figure 7, are required to play the role of launching the correlations; farming out the pieces of the correlations to the compute nodes; collecting and combining the processed correlation pieces and writing out the correlated data products. These 4U 2 processor systems with 24 3.5" drive headend nodes utilize the onboard hardware SAS RAID controller to achieve high output data rates and data protection.

Figure 7. *Supermicro 4U Headend Node Systems (SSG-6048R-E1CR24H)*

There are a total of 60 compute nodes that comprise the MIT cluster. There are 38 nodes of the Supermicro TwinPro multi-node systems (19x systems total), shown in Figure 8 with Intel Xeon E5-2640 v4 processors. The Twin multi-node system contains two independent dual processor compute nodes in a single system doubling the density from traditional rackmount systems and with shared power and cooling for improved power efficiency and serviceability.

Figure 8. *Supermicro TwinPro Multi Node Systems (6028TP-DNCTR)*

There are 16 previous generation compute cluster nodes in the cluster comprised of a 3U dual socket Supermicro server with the Intel Xeon E5-2680 v2 processors, Figure 9.

Figure 9. *Supermicro SuperServer (SYS-6037R-72RFT)*

The clustered storage nodes are configured with redundant high efficiency power supplies and optimized redundant cooling to save energy, SAS3 expander options for ease of interconnection, plus a variety of drive bay options depending on task.

At the core of the MIT DiFX Correlator is a high-performance data storage cluster based on four Supermicro storage systems, to deliver high I/O throughput and data availability through 10 Gigabit Ethernet networking fabrics and RAID controllers.

These systems are built with a selection of different Supermicro serverboards and chassis with support for dual or single Intel® Xeon® processors, SAS3 drives with onboard hardware RAID controllers, onboard dual 10GbE for efficient networking, up to 2 TB DDR4 memory and 7 PCI-E 3.0 expansion slots for external drive capabilities.

Processing the EHT Data

According to Vincent Fish2, "The time for computation is in general a complicated function a lot of different parameters (not just the number of stations or baselines being correlated). In general, we can correlate the data from one 2 GHz chunk of EHT data a bit slower than real time. However, the telescopes don't record continuously—there are gaps between scans on different sources, and an observing night doesn't last 24 hours—so we could correlate a day's worth of data in about a day per 2 GHz band if we ran the correlator continuously.

The 2017 data consisted of two 2 GHz bands, one of which was correlated at Haystack and the other in parallel at MPIfR (MPI). The 2018 data consists of four 2 GHz bands; each correlator is responsible for two of them."

Figure 10. *Mark 6 Playback Unit*

The Mark 6 playback units, figure 10, at the MIT correlator are connected via 40 Gbps data links. A 100 Gbps network switch then delivers data to the processing nodes using 25 Gbps links. At MPI the internode communication, which includes the Mark 6 playback units, is realized via 56 Gbps connections, exceeding the maximum playback rate of the Mark 6 units of 16 Gbps3.

The average time and bandwidth in the correlators are set to ensure that any coherence losses due to delay or rate variations are negligible, or equivalently that such variations can be tracked both in time and frequency.

The processing was divided between the two sites and included crosschecking of results. The supercomputers at MPI and MIT correlated and processed the raw radio telescope data from the various observing sites.

After the initial correlation, the data are further processed through a pipeline that results in final data products for use in imaging, time-domain analyses, and modeling.

Data were correlated with an accumulation period (AP) of 0.4 s and a frequency resolution of 0.5 MHz.

Note that ALMA refers to the Atacama Large Millimeter/submillimeter Array (In Chile). ALMA was configured as a phased array of radio telescopes and was a recent addition to the EHT effort with significant resolution capability. ALMA was treated as a highly accurate anchor station and thus was used to improve the sensitivity limits of the global EHT array.

Although operating as a single instrument spanning the globe, the EHT remains a mixture of new and well-exercised stations, single-dish telescopes, and phased arrays with varying designs and operations. Each observing cycle over the last several years was accompanied by the introduction of new telescopes to the array, and/or significant changes and upgrades to existing stations, data acquisition hardware, and recorded bandwidth.

EHT observations result in data spanning a wide range of signal-to-noise ratio (S/N) due to the heterogeneous nature of the array, and the high observing frequency produced data that were particularly sensitive to systematics in the signal chain. These factors, along with the typical challenges associated with VLBI, motivated the development of specialized processing and calibration techniques.

The end result of all this work, involving an intense international collaboration, was the first image of a black hole event horizon.

The Future of Black Hole Observations

The EHT team's initial goal was to image the event horizon of the massive black hole at the center of our own Milky Way galaxy, SgrA*. However, this proved more difficult than originally anticipated, since its structure changes on the timescale of minutes.

According to Max Planck Director, Anton Zensus[5], "the heart of our Milky Way is hidden in a dense fog of charged particles. This leads to a flickering of the radio radiation and thus to blurred images of the center of the Milky Way, which makes the measurements more difficult. But I am confident that we will ultimately overcome this difficulty. On the other hand, M87 is about 2,000 times further away. However, the black hole in its center is also about 1,000 times more massive than the one in our Milky Way. The greater mass makes up for the greater distance. The shadow of the black hole in M87 therefore appears to us to be about half the size of the one from the gravity trap in our Milky Way."

To date, the EHT has observed the black holes in just one wavelength—light with a wavelength of 1.3 millimeters. But the project soon plans to look at the 0.87-mm wavelength as well, which should improve the angular resolution of the array. It should also be possible to sharpen the existing images using additional algorithmic processing. As a consequence, we should expect better images of M87 and other black holes in the not too distant future. The addition of more participating radio telescope sites will also help improve the observational imaging.

The EHT team also wants to move from only ground-based VLBI to space-based imaging using a space-based radio telescope. Going into space would allow the EHT to have radio telescopes that are even further apart and thus able to capture some even more astounding and higher resolution images of the black holes around us. "We could make movies instead of pictures," EHT Director Sheperd Doeleman said in an EHT talk at the South by Southwest (SXSW) festival in Austin, Texas[5]. "We want to make a movie in real time of things orbiting around the black hole. That's what we want to do over the next decade."

[5] "An astounding coincidence with theory", Anton Zensus, Max Planck Director interview, https://www.mpg.de/13326245/interview-anton-zensus-max-planck-eht

One of the big obstacles to using a space based EHT dish is data transmission. For the ground-based experiments, HDDs were physically transported from the telescope sites to central processing facilities at MPI and MIT. It is not clear yet how data would be sent from the space telescopes to earth, but laser communication links are one possibility. Transferring large amounts of data to the ground may require substantial onboard data storage and a geosynchronous satellite acting as a relay (Figure 6).

Deepening our understanding the universe around us requires sophisticated IT infrastructure. This includes ever more digital processing with more advanced algorithms, using faster and more sophisticated servers and fast as well as vast digital storage for capturing and processing the sea of data generated by big international science projects, such as the Event Horizon Telescope project.

Infrastructure Possibilities of Future Black Hole Observations

Future work to gain higher resolution and even time sequenced data (e.g. videos) of black hole event horizons (including the black hole at the center of our Milky Way galaxy) will involve new data and more sophisticated analysis algorithms as well as the use of radio telescopes in space. These efforts can leverage the massive improvements already available in today's state-of-the-art IT Infrastructure.

The core count of the 10 Twin systems used in the current correlator could be achieved with a single Supermicro BigTwin™ multi-node system with 2U 4 Nodes, Dual Socket 205W Intel Xeon Scalable Processors, 24 DIMMS DDR4 memory with 6 All-Flash NVMe drives per node. The system delivers better density and improved power efficiency.

Figure 11. *Supermicro BigTwin™*

---

[5] Effort to 5 SXSW 2019 Panel, Event Horizon Telescope: A Planetary Photograph a Black Hole, Sheperd Doeleman, Dimitrios Psaltis, Sera Markoff, Peter Galison, https://eventhorizontelescope.org/blog/eht-planetary-effort-photograph-black-hole

The rendering of images could be accelerated from hours to seconds with advanced GPU systems such as a 1U 4-GPU Server, Figure 12.

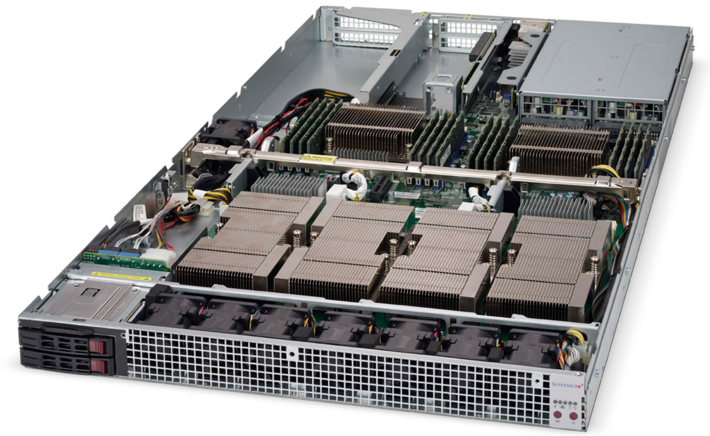


Figure 12. *Supermicro GPU-optimized server systems*

The 960 terabytes of data could be stored on a single 1U Petascale server with the All-Flash NVMe Solid State drives with order of magnitude better performance, reduced latency and eliminating environmental issues introduced from the high altitude, Figure 13.

Figure 13. *Supermicro All-Flash NVMe storage systems*

These are just a few examples of the new state-of-the-art IT Infrastructure available to researchers across the globe to support and enhance future research and discovery.

About the Author

Tom Coughlin, President, Coughlin Associates is a digital storage analyst as well as a business and technology consultant. He has over 37 years in the data storage industry with engineering and management positions at several companies.

Dr. Coughlin has many publications and six patents to his credit. Tom is also the author of Digital Storage in Consumer Electronics: The Essential Guide, which is now in its second edition with Springer. Coughlin Associates provides market and technology analysis as well as Data Storage Technical and Business Consulting services. Tom publishes the Digital Storage Technology Newsletter, the Media and Entertainment Storage Report, the Emerging Non-Volatile Memory Report and other industry reports. Tom is also a regular contributor on digital storage for Forbes.com and other blogs.

Tom is active with SMPTE (Journal article writer and Conference Program Committee), SNIA (including a founder of the SNIA SSSI), the IEEE, (he is past Chair of the IEEE Public Visibility Committee, Past Director for IEEE Region 6, President of IEEE USA and active in the Consumer Electronics Society) and other professional organizations. Tom is the founder and organizer of the Storage Visions Conference (www.storagevisions.com as well as the Creative Storage Conference (www.creativestorage.org). He was the general chairman of the annual Flash Memory Summit for 10 years. He is a Fellow of the IEEE and a member of the Consultants Network of Silicon Valley (CNSV). For more information on Tom Coughlin and his publications and activities go to www.tomcoughlin.com.

About Super Micro Computer, Inc.

Supermicro® (NASDAQ: SMCI), the leading innovator in high-performance, high-efficiency server technology is a premier provider of advanced server Building Block Solutions® for Data Center, Cloud Computing, Enterprise IT, Hadoop/Big Data, HPC and Embedded Systems worldwide. Supermicro is committed to protecting the environment through its "We Keep IT Green®" initiative and provides customers with the most energy-efficient, environmentally-friendly solutions available on the market.

www.supermicro.com

Printed in USA Please Recycle

Super Micro Computer, Inc.
980 Rock Avenue
San Jose, CA 95131 USA
www.supermicro.com