Distributed Processing Goes Galactic

by H. Paul Shuch, N6TX
Executive Director, The SETI League Inc.

Abstract

Are we alone, the sole sentient species in the cosmos, or might there be others, among whom we can take our rightful place? If there is indeed an interstellar internet, might we someday log on? And what are the protocols for cosmic communication? For the first time in human history we now have the technology to ask, and perhaps begin to answer, these questions. In this paper we explore the strengths and weaknesses of SETI@home, the most ambitious distributed computing experiment on this planet. We learn how thousands of amateur radio telescopes are forming a global net to snare that elusive fish in the cosmic pond. And we explore how the lessons learned from the SETI@home experience can be brought to bear on the problem of massive data collection and analysis. The author believes it is the world's radio amateurs and computer hobbyists who will ultimately bring in signals from the stars.

SETI 101: An Introductory Course

SETI, the electromagnetic Search for Extra-Terrestrial Intelligence, is a relatively young science with a colorful history, which seeks to detect direct radio evidence of other technological civilizations in the cosmos. For forty years, its dominant paradigm has been the use of giant radio telescopes using sensitive microwave receivers and powerful computers, to scan nearby stars for artificially generated signals of intelligent alien origin. Once funded through NASA, SETI research in the United States lost its government support seven years ago, and now continues as a privatized venture conducted by various grass-roots nonprofit organizations.

Giant radio telescopes (such as the 1000-foot spherical reflector of the Arecibo radio observatory in Puerto Rico) achieve part of their sensitivity by directing an extremely narrow beam on the heavens. Such instruments view perhaps one millionth of the sky at a given time, reducing the received background noise (and hence improving the signal to noise ratio of any detected radio artifact) by a factor of a million, relative to an omnidirectional (isotropic) antenna. Hence, if you have such an antenna hooked up to the right kind of receiver, which is tuned to precisely the right frequency, at exactly the instant when The Call comes in, there is about a 99.9999% chance it will be pointed the wrong way.

But since we don't know what exactly that "right frequency" is, the problem of SETI success is complicated further by our need to tune our receivers systematically across a wide spectral range. If the a SETI receiver is to achieve reasonable sensitivity, its desired bandwidth of reception is of necessity quite narrow. This is because radio noise and natural interference phenomena are spectrally broad, while one of the hallmarks of artificiality which intelligently generated emissions can be expected to demonstrate is high spectral coherence, resulting in narrow bandwidth (on the order of a fraction of a Hertz.) But that narrow signal can fall anywhere within several GHz of potential spectral real estate.

By reducing receiver bandwidth, we once again improve our signal to noise ratio. But if we wish to find an artificial signal with a bandwidth of, say, 1 Hz, and if to locate it we must scan, for example, 10 GHz of spectrum, then we have a temporal problem equivalent to our previously stated spatial one. If we are pointed in exactly the right direction, at the very instant when The Call comes in, there is something like a 99.99999999% chance we'll be tuned to the wrong frequency!

Problems associated with the spatial dimension of SETI are partly overcome by constructing phased arrays of a great many antennas, operated so as to look in many directions at once. And to address the spectral challenge, SETI engineers concentrate much of their efforts toward developing elaborate multi-channel spectrum analyzers (MCSAs) capable of monitoring many millions of narrow channels simultaneously. Yet despite these technological advances, SETI critics rightly point out, after forty years of SETI, we have yet to detect a single confirmed signal of intelligent extraterrestrial origin.

The SETI@home Experience

Yet in those forty years of searching, SETI proponents counter, we have not only failed to scratch the surface, we haven't even felt the itch. Our massive antennas and multi-channel spectrum analyzers generate more data than we can ever hope to analyze, even using Planet Earth's most power supercomputers. Digital signal processing efforts, necessary to separate the cosmic wheat from the galactic chaff, depend upon an ability to crunch numbers at an ever increasing rate. And right now we are generating more SETI data than we can ever hope to analyze.

It has been argued that if we are to be limited at all, computing power is the place to be limited. Since computer power seems to double every year or so, we need merely wait until our computer power is up to the task, in just a few more years -- or decades -- or centuries.

One group of SETIzens, lead by Prof. Woody Sullivan at the University of Washington, and Dan Werthimer, David Anderson and David Geddey at the University of California, Berkeley, got tired of waiting. Last year they launched SETI@home, planet Earth's most ambitious distributed processing experiment. They have harnessed the idle processing power of 1.5 million personal computers around the world, and in so doing have created the world's most powerful supercomputer. Tied to a Project SERENDIP microwave receiver at the Arecibo Observatory, the SETI@home network has crunched more data in a few short months than has been analyzed by all the world's prior SETI efforts, combined.

After pre-processing on site, SERENDIP cranks out data at the rate of more than 3 MBytes per second. That's just under 500 CD-ROMs per day worth of data generated by one receiver alone, truly an example of drinking from the fire hose. Until now, most of that data would simply have gone un-analyzed.

SETI@home's power derives from clever software that parses this massive data stream into bite-sized chunks, for Internet distribution and off-line data analysis. This is done by first filtering and dividing the receiver's 10 MHz bandwidth into 1024 individual audio channels, each a manageable (and easily digitized) 9756.6 Hz wide. 107 seconds worth of data from just one of these audio channels can be stored in a digital file a mere 341 kBytes in size (about the capacity of a 1980s-vintage single sided, single density 5 1/4 inch floppy disk).

At typical V.90 dial-up connect speeds, it takes about a minute and a half to transfer such a packet to a SETI@home user via the Internet, after which a modest computer running the SETI@home analysis software can thoroughly analyze it offline. On a typical Pentium III class PC, this analysis takes roughly ten hours, after which the user uploads a results file to the SETI@home server at Berkeley, downloads another packet, and starts all over again.

Working together is certainly working! Today one and a half million home computers are devouring data from the world's largest radio telescope, TeraBytes at a time. Still, while the screen saver churns away in the background, the appetite for involvement is not sated. "I'm no rocket scientist," I hear you saying, "but I want to do more than wait for my Pentium to claim the prize. Where can I go from here? The software is fully capable of discovering that elusive needle. Only, where do we find the haystack?"

SETI@home has some impressive strengths, offset by one significant weakness. Its 1.5 million home computers are crunching data from a single source, an antenna which in turn sees only a tiny fraction of the sky at any time. To avoid missing the call, we really need about a million such radio telescopes, coordinated so as to point in all directions at once. But at a cost of perhaps a hundred million dollars apiece, we've just exceeded the Gross Planetary Product.

Fortunately, there is another way.

Project ARGUS and the Amateur Radio Astronomer

Launched in 1996, Project Argus is an amateur-run all-sky survey which attempts to accomplish something which NASA SETI never contemplated: see in all directions at once. This major initiative of the membership-supported, nonprofit SETI League seeks to harness the power of 5000 small radio telescopes worldwide, in a coordinated search of all four pi steradians of space. Its amateur radio telescopes, typically built around discarded satellite TV antennas for a few hundred to a few thousand US dollars, achieve sensitivities on the order of 10^-23 watts per square meter, roughly equivalent to the best research-grade radio telescopes of the late 1970s. As personal computers and digital signal processing software become more powerful, this two-decades gap between professional and amateur capabilities is beginning to narrow.

One argument for the validity of Project Argus is the example set by amateur optical astronomers in their discovery of numerous comets, supernovae, and other highly intermittent astrophysical phenomena. It is not, after all, the world's great observatories which typically detect these events, but rather such dedicated amateur astronomers as Allen Hale and David Levy (both using 14 inch Schmidt Cassegrain telescopes), Tom Bopp (who doesn't even own a telescope, but discovered his comet with one he borrowed from his astronomy club), Hyakutaki (the comet that bears his name was discovered with a pair of high-power field binoculars!) and the late Gene Shoemaker, a geologist by trade, but a longtime and avid amateur skygazer.

Regrettably, the analogy breaks down when one considers equipment availability. In most cities of the world an aspiring comet-hunter can walk into a local optical shop, write a check for a thousand dollars or so, and walk out with a telescope that would have turned Galileo green with envy. Amateur radio astronomers aren't quite so fortunate. You can't walk in to your local Radio Shack ® store and buy a radio telescope. At least, not yet. The SETI League is trying to change that, by designing the hardware (licensed for commercial manufacture) and software (distributed as shareware via the internet) to turn a surplus 3 to 5 meter TVRO dish and a home computer into a credible research instrument. Already about 100 radio amateurs, microwave hobbyists, electronics experimenters and computer hackers around the world have succeeded in putting their Argus stations on the air. Hundreds more are now under construction, and every year the dream of all-sky coverage (whereby no direction on the sky shall evade our gaze) comes closer to reality.

Details on the construction of a Project Argus radio telescope may be found in the SETI League Technical Manual, available in hardcopy for a modest contribution, or free on the web (www.setileague.org).

Global and Galactic: The ARGUS@home concept

Current Project Argus instruments each scan about 22 kHz of frequency spectrum at a time (a small fraction of the 10 MHz instantaneous bandwidth of the SERNENDIP receiver at Arecibo), and break that down into typically 8192 simultaneous channels, each about 2 1/2 Hz wide. One such instrument generates on the order of 44 kBytes per second of data. This represents but 1.5% of the data gathered by the SETI@home experiment at Arecibo. On the other hand, the existing 100 Argus stations, collectively, already exceed the data output of a SERENDIP receiver. By the time Argus reaches full strength, its combined network of 5000 amateur radio telescopes will generate as much data as 70 Arecibos!

The SETI@home packet your PC is processing came from the world's largest radio dish. So did everybody else's. Which means that a million and a half PCs are being serviced by a single data source. A powerful source to be sure. But with lotteries all over the world, why buy all our tickets for a single drawing? Remember that Arecibo achieves its sensitivity by scanning a slim slice of the celestial sphere. No software in the world is going to find photons that didn't hit the fan. No matter how many computers are running it.

Perhaps that's where the eyes of Argus can really shine. Imagine a global network of thousands of amateur radio telescopes, scanning the entire sky in real time. Now imagine something akin to SETI@home, software which will let you scan that data via the Internet. Only instead of archival data recorded weeks ago, we're talking live data which your computer can capture in real time. So you need not wait for the evening news to hear the winning numbers.

ARGUS@home won't happen overnight, any more than SETI@home did. In addition to the multitude of small radio telescopes required, we still need to come up with a SETI@home-compatible data block format, and find a way for Project Argus software to parse out the gathered data for internet distribution. Then there's the challenge of collecting and correlating all those processed packets. The SETI@home experiment has already solved many of these problems; it remains for The SETI League to adapt their solutions to amateur practice. We hope that by the time SETI@home drinks the Arecibo well dry, we will have risen to these technical challenges.

Conclusion

Project Argus went online just four years ago with only five small amateur radio telescopes. Today we're running about a hundred. It's going to take us a few more years before the Argus network grows to truly global proportions. Until then, there's always Arecibo.

The distributed computing concept pioneered by SETI@home is very adept at finding needles. The global network of Argus telescopes will be ideal for finding haystacks. Seems to me, it's a marriage made in heaven.

About the Author

Dr. SETI ®, as H. Paul Shuch is known to his intimates, is something of a cross between Carl Sagan and Tom Lehrer (he sings like Sagan and lectures like Lehrer). The aerospace engineer credited with the design of the world's first commercial home satellite TV receiver, Paul now directs his microwave interests toward the search for life in space. Dr. Shuch received his Ph.D. in Engineering from the University of California, Berkeley, and taught for 24 years. He is the founding Executive Director of The SETI League, Inc., a membership-supported educational and scientific nonprofit which has emerged as the leader in a grass-roots Search for Extra-Terrestrial Intelligence.

As a radio amateur, N6TX has been active on all 20 ham bands between 1.8 MHz and 24 GHz. Paul is the author of over 250 publications. He is a Fellow of the British Interplanetary Society, serves as a fellowship interviewer for the Hertz Foundation, a manuscript reviewer for several peer reviewed journals, has been an advisor to the National Science Foundation, and is a military program evaluator for the American Council on Education. Paul's honors include the Robert Goddard Scholarship, the Hertz Fellowship in the Applied Physical Sciences, the Horonhjeff Grant, the Hertz Doctoral Thesis Prize, the EAA Safety Achievement Award, the John Chambers Memorial Award, the ARRL Technical Achievement Award, and the Dayton Hamvention Technical Excellence Award.

Dr. SETI lives on a radio-quiet hilltop north of Williamsport PA (grid square FN11lh) with three radio telescopes, his wife Muriel Hykes, and five of their seven recombinant DNA experiments.

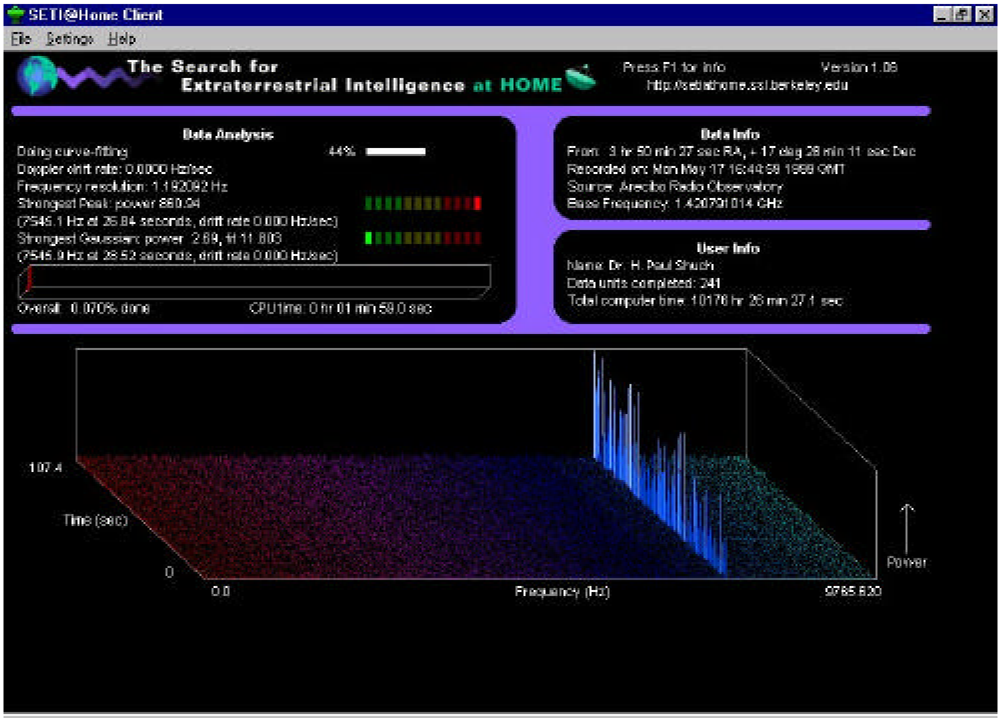

Figure 1: Strong, coherent signals such as this one quicken the pulse of many a SETI@home project participant. Unfortunately, all so far have been generated not by ETI, but rather by terrestrial inteference, or by the Wizards of Arecibo as they inject test signals to verify the proper operation of their equipment.

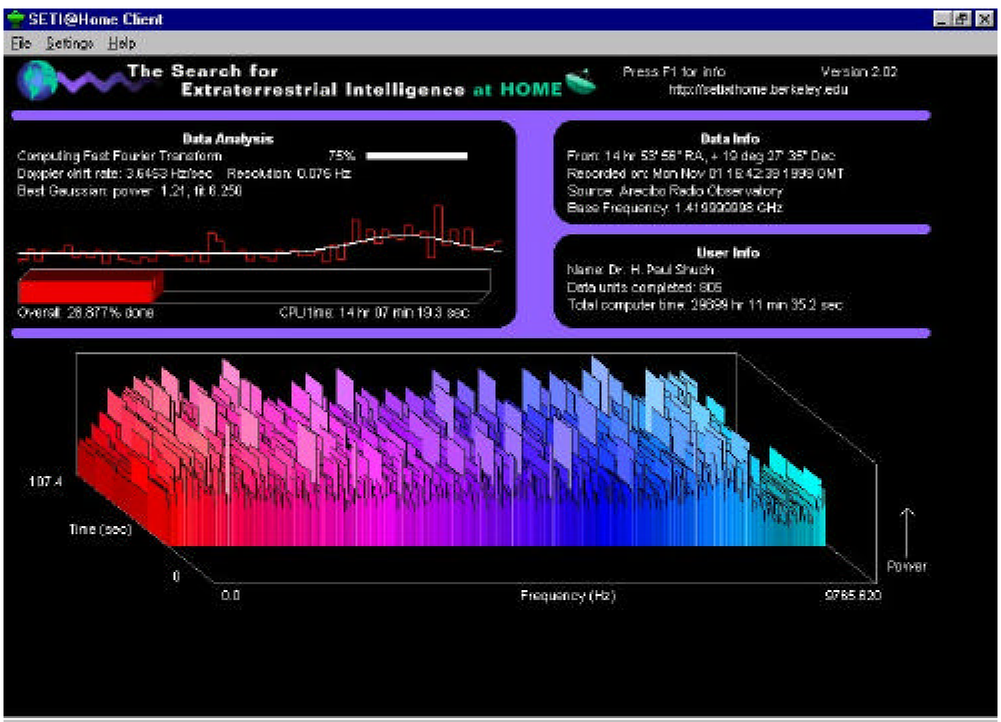

Figure 2: The SERENDIP receiver at Arecibo has a 10 MHz instantaneous bandwidth. For SETI@home processing, its output is parsed out into 1024 sub-bands each 9765 Hz wide. Notice how the amplitude of the noise rolls off at both the top and the bottom frequency ends of this analysis spectrogram. The curve seems to follow the frequency response of an audio bandpass filter optimized to the desired sub-channel bandwidth.

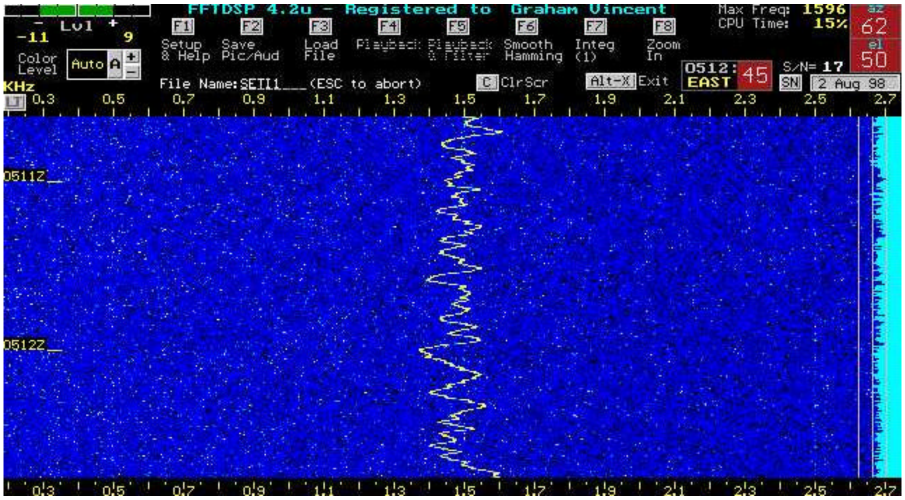

Figure 3: Graham Vincent, SETI League volunteer coordinator for New Zealand, received this intriguing signal on 2 August 1998, at a frequency of 1281.919 MHz USB (in the 23 cm amateur radio band). The appearance of the signal is similar to a class of anomalies detected by the SETI Institute's Project Phoenix targeted search. Dubbed "wigglers" by the SETI Institute's Dr. Jill Tarter, their detections have always proved to be cases of radio frequency interference. Graham's signal is no exception. It turned out to be computer rfi, generated within the very computer which he was using to run his signal analysis software. England co-coordinator Ken Chattenton, who has had similar experiences, recommends that if a signal is strong enough to be audible, one should turn off the computer and see if it goes away.

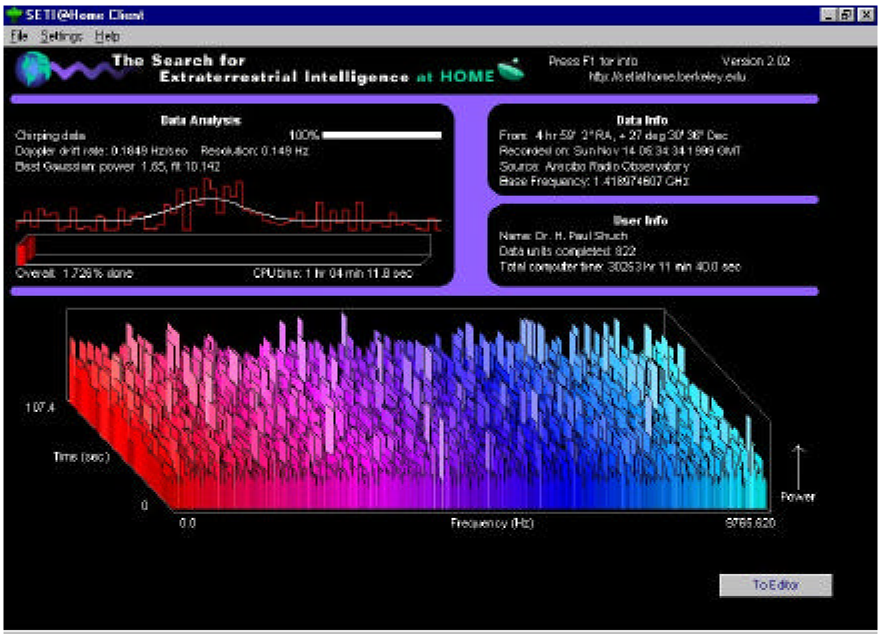

Figure 4: Here's a signal with a fairly good Gaussian fit, which is not evident from viewing just the 3D spectrogram (bottom window). The SETI@home client divides the 9765 Hz wide data block into housands of very narrow bins. The amplitude of the signal in each individual bin is analyzed over time, and the bin with the best fit to the antenna's expected drift-scan time series is displayed as a ragged trace in the data Analysis (upper left hand) window. The smooth trace represents an ideal Gaussian curve (normal Distribution) corresponding to the pattern of the Arecibo antenna. The two curves are statistically compared. The closer the fit, the more credence is given a candidate signal. Of course, the Gaussian test is only one of many hurdles a signal must pass before it is considered to be of extra-terrestrial origin.

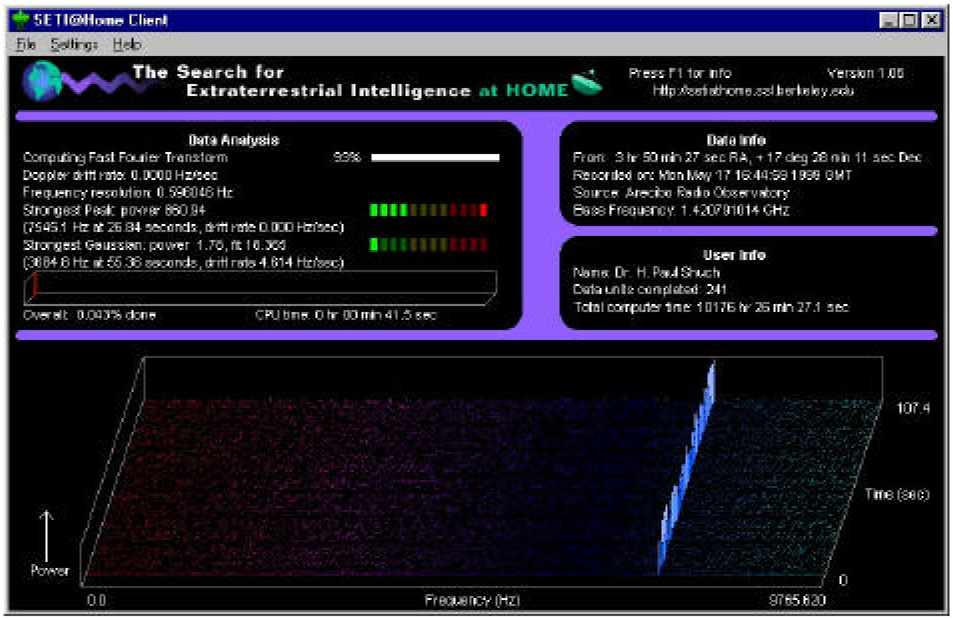

Figure 5: SETI@home's 1.5 million users continue to see occasional anomalies such as this one, observed by the author in December 1999. Members sometimes call or email The SETI League, requesting that we check out such signals (most of which turn out to be terrestrial interference). Unfortunately, there's nothing we can do from here to analyze these detections, since all verification is performed by the SETI@home team at Berkeley CA. Be sure to uploaded your analysis files to them, and rest assured that they will indeed follow-up on all interesting candidatesignals, and inform you if yours is The One.

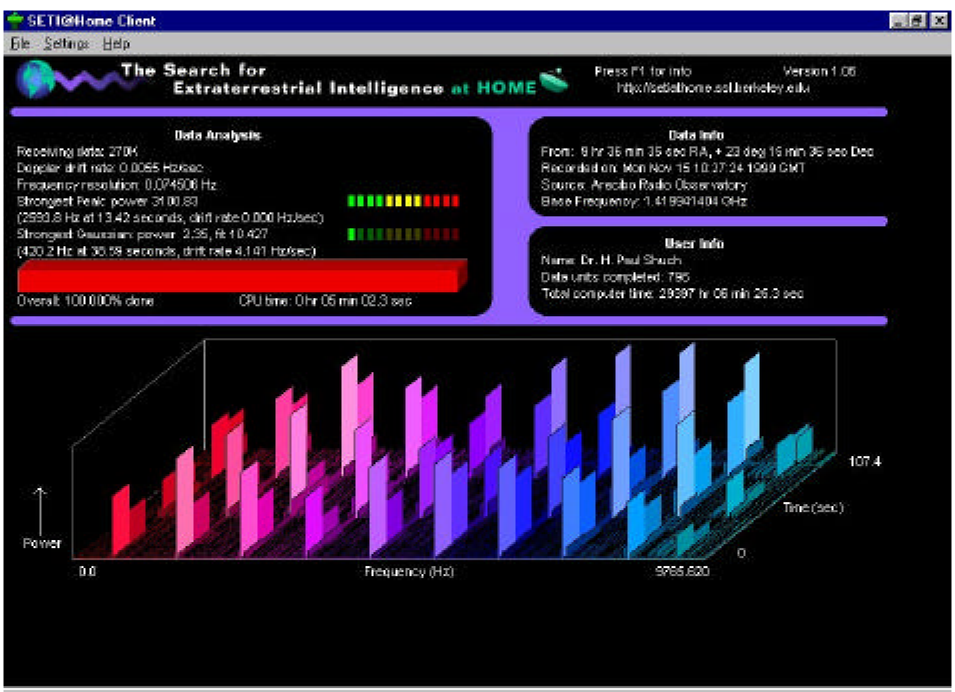

Figure 6: The typical computer takes tens of hours to fully analyze a single SETI@home data block. Occasionally, strong, wideband terrestrial interference obliterates any useful information. When that happens, the SETI@home clinet determines that no further analysis of that data block is possible, quickly terminates analysis of that particular file, and requests another one for analysis. SETI@home gave up on this file after just five minutes of attempted analysis.

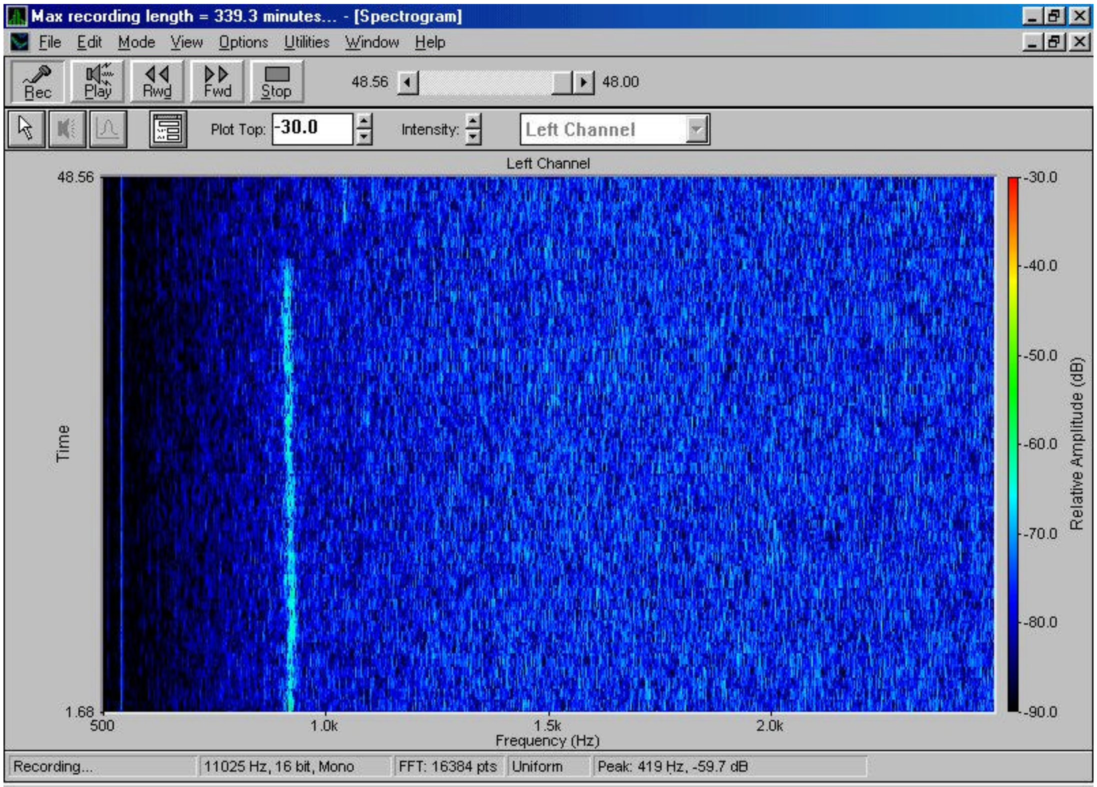

Figure 7: EME (moonbounce) contests provide Project Argus participants with an opportunity to detect weak amateur microwave signals reflected off the lunar surface. This unusually strong 1296.015 MHz EME echo from the 30 foot dish of Jay Leibmann, K5JL, was received at Argus station FL11LH during the 30 October 1999 ARRL EME contest.

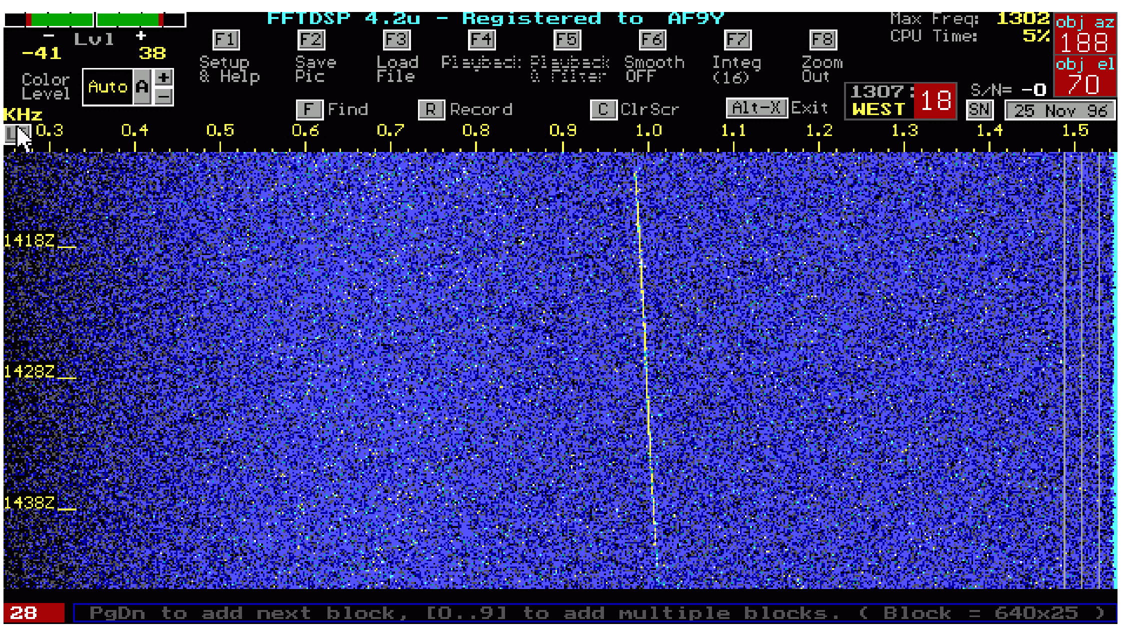

Figure 8: This CW signal from the Mars Global Surveyor was received by SETI League member Mike Cook on 25 November 1996, while the spacecraft was about 5 million km from Earth. The satellite's 1.3 Watt beacon transmitter, into an omnidirectional antenna, provided SETI enthusiasts with an excellent dry run to verify the operation of their receivers and digital signal processing software. Several other SETI League members were also able to recover the signal utilizing Mike's FFTDSP shareware program.