THE CLIPPER GROUP
Navigator™
Navigating Information Technology Horizons℠
Published Since 1993 Report #TCG2006108 December 20, 2006

Code Green Networks
Catches Information Leakage as It Happens

Analyst: Anne S. MacFarland

Management Summary

Whether a business is large or small, opportunities for information leakage continue to grow. Laptops are taken everywhere, networks are used that are not fully secure, and removable storage devices such as thumb drives give more and more convenient capacity at less and less cost. When you add in search and data classification, which make valuable business information easier to find, you have a perilous situation.

Documentation of business experience – good or bad – has always been important. In electronic form, it can be analyzed quickly in far more detail than ever before to pinpoint trends and opportunities that would otherwise be hard to find. While buzz and hype may be satisfying to the casual user, it is the R&D document and the business records that have high business value for competitors with the intelligence to analyze them. Improper exposure of information is a huge business risk. To top off this parfait of risk, disclosure of information – customer information, intellectual property or any other information that may impair the business must be reported by many organizations, such as companies that are publicly traded. Any organization that partners, even if not mandated to report by regulations, will find reporting a breach to these partners is an essential part of business etiquette.

These risks are not adequately met by *ex post facto* filtering or routing documentation. They are not addressed by archival arrays, which preserve the evidence but do nothing to prevent or even identify the event. T**he good news is that there are things that can be done, at a price reasonable to most enterprises, to mitigate these information leakage risks.** A company called Code Green Networks, based in Santa Clara, California, has just released an appliance that does content filtering for corporate networks and outbound and inbound electronic communications. Code Green brings a wealth of intellectual property to bear in its solution, from the basic algorithms needed to prevent the exposure of social security and account numbers to what it calls and has patented as *Deep Content Fingerprinting*. This Deep Content Fingerprinting allows a phrase, sentence or paragraph to be identified whenever it is transmitted – in any format – in a derivative document – and even if it is SSL encrypted.

This changes everything. It changes the definition of the "best efforts" to which SOX-compliant enterprises must strive. It gives a control system for the information itself, not just the data files and document chunks that are classified and tagged by content and document management systems or file systems. And it gives enterprises new ways to reduce the ungainliness of the risk profile that is an inevitable part of doing business.

In addition, while the solution scales large, the Code Green Networks appliance is targeted at the mid sized enterprise and priced accordingly. The risk mitigation they offer is something that should be of interest to any business. To learn more about how they do it, please read on.

IN THIS ISSUE

The Issue of Information Leakage

There are different kinds of information exposures.

- **It may be accidental**, as when someone hits send before checking that the auto-complete of the sender address was accurate. It may be inadvertent, due to sloppy maintenance of group mail lists, forgetting to erase the correspondence trail when forwarding a document.
- **It often is a regular part of doing business.** Increasingly, such correspondence is encrypted. Enterprises still need to watch the transmission of sensitive information to see that the patterns conform to normal use.
- **It can also be intentional and malicious** – perhaps not for monetary gain but to curry favor, or build relationships. That no money is involved does not make it benign.

Code Green's Approach

Most enterprises have exposures in all three areas. To address these risks, unacceptable disclosure of information must be identified when it happens, every time it happens. Code Green addresses the problem by providing the following.

Appliance

The *Code Green Networks Content Inspector*, *CI 1500* (for specifics, see Exhibit 1 at right), is usually deployed as an out-of-band device. It can be paired for high availability, and ganged for greater scale. For outgoing e-mail, Code Green recommends connecting the appliance(s) to the email server as proxies via a SNMP MPA connector. The appliance can also be used to track sensitive information, bad language, etc on enterprise networks.

Business Rules

The CI 1500 comes with pre-loaded business rules to cover basic business functions. These include rules to identify transmission of standard obscenities or social security and national ID numbers. Standard rules for particular industries, such as health care, are developed and evolved by industry organizations, and can be added. With these rules, the appliance can stop most of the information leakage. Corporate risk assessors and company lawyers will have a good idea of what additional rules are needed for a specific enterprise.

Code Green CI 1500

- Dual core 64-bit Xeon processors, dual fans, dual power supplies
- 6 GbE ports
- Ruggedized Fedora Linux OS.
- RAID 5 storage
- Supports Fingerprints on up to 1 TB of source content.
- Handles unstructured and structured content via pattern matching.
- Can analyze content in .zip and .tar files, even nested files.
- Can process half a million emails/hour with up to 200 Mbps of throughput.
- *Source: Code Green Networks*

Filters

To business use, a range of filter options is needed. For many exposures, keywords (taxonomies, etc) address the risk, along with timely updating to include, say, the names of new acquisition targets. Pattern-matching addresses exposure of certain kinds of files, product information, etc. But for information-rich businesses like legal firms, a larger granularity of filter that can catch particular phrases, sentences, or even paragraphs is useful - and it is useful for more than just preventing information leakage.[1]

Code Green can do this larger-scale information control in a patented method it calls this Deep Content Fingerprinting. This involves taking a word, phrase or paragraph, converting it to Unicode, tokenizing, hashing it and winnowing out information so that the fingerprint cannot be rebuilt (for security reasons).

Actions

Again, any enterprise will want a range of action options. Some kinds of information exposure merit a logging and/or e-mail notice. Others demand a block-and-quarantine. Other situations are best addressed by rerouting the information to an Email encryption server. Code Green has partnerships with two encryption vendors.

[1] For any enterprise dealing with contracts or negotiations, the ability to check consistency of clauses is a huge deal. For any company dealing in source code, the ability to track this code is huge.

The Code Green approach is not like the membership approach of Digital Rights Management or Content Management applications, where the source content must be selected, ingested, and processed before it can be controlled. Such alternatives usually control a limited domain of content, and the information controlled (at various levels of granularity) is festooned with metadata that can bulk up the storage needed to persist it. With Code Green, all data is addressed – but it is crawled and compared only when it is transmitted – because that is when the large business risk occurs.

Who Is Code Green Networks?

Code Green Networks was founded in 2004 by Sreekanth Ravi and Sudhakar Ravi, brothers who previously founded Sonic Wall. The product they have built has been in an extensive beta with hundreds of users in many geographies using many languages. Their beta customers (who, for obvious reasons, prefer not to be identified), span the expected industries of financial services, legal, government, manufacturing, and technology.

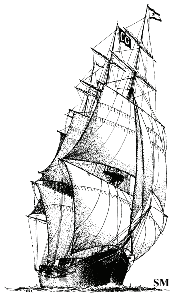
Code Green distributes its product through its value-added reseller channel in the United States, Europe, and Japan. Pricing of the appliance is $25,000 for up to 250 users, $50,000 for up to 1000 users, and $100,000 for unlimited use. Code Green also offers a leasing option of $500/appliance/month, which will be an attractive way for many smaller enterprises to try before they buy.

Once a Sales Engineer deploys a Code Green appliance on the network (in read-only mode, so it is not a security threat), in a few hours it can report on the extent of enterprise data leakage. If leakage is negligible, the appliance can be removed. However, this seldom happens.

Code Green has an aggressive roadmap of future enhancements. It is developing connectors for databases, similar to those it has for EMC *Documentum* and Oracle *Stellent* content management products. It is considering how best to add discovery to aid businesses in finding sensitive content that needs to be protected.

Conclusion

Even without the goals of mandated best efforts and reporting of data exposures, the value of your information and the opportunism of your competition demands that you address this problem. The range of options that Code Green provides will help avoid corporate risk, and at the same time expose many ways to use corporate information better.



About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

***Anne MacFarland* is Director of Data Strategies and Information Solutions for The Clipper Group.** Ms. MacFarland specializes in strategic business solutions offered by enterprise systems, software, and storage vendors, in trends in enterprise systems and networks, and in explaining these trends and the underlying technologies in simple business terms. She joined The Clipper Group after a long career in library systems, business archives, consulting, research, and freelance writing. Ms. MacFarland earned a Bachelor of Arts degree from Cornell University, where she was a College Scholar, and a Masters of Library Science from Southern Connecticut State University.

- ***Reach Anne MacFarland via e-mail at Anne.MacFarland@clipper.com or at 781-235-0085 Ext. 128. (Please dial "128" when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

The Clipper Group Navigator, **The Clipper Group Explorer**, **The Clipper Group Observer**, **The Clipper Group *Captain's Log***, **The Clipper Group Voyager**, and *"clipper.com"* are trademarks of The Clipper Group, Inc., and the clipper ship drawings, *"Navigating Information Technology Horizons"*, and *"teraproductivity"* are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.