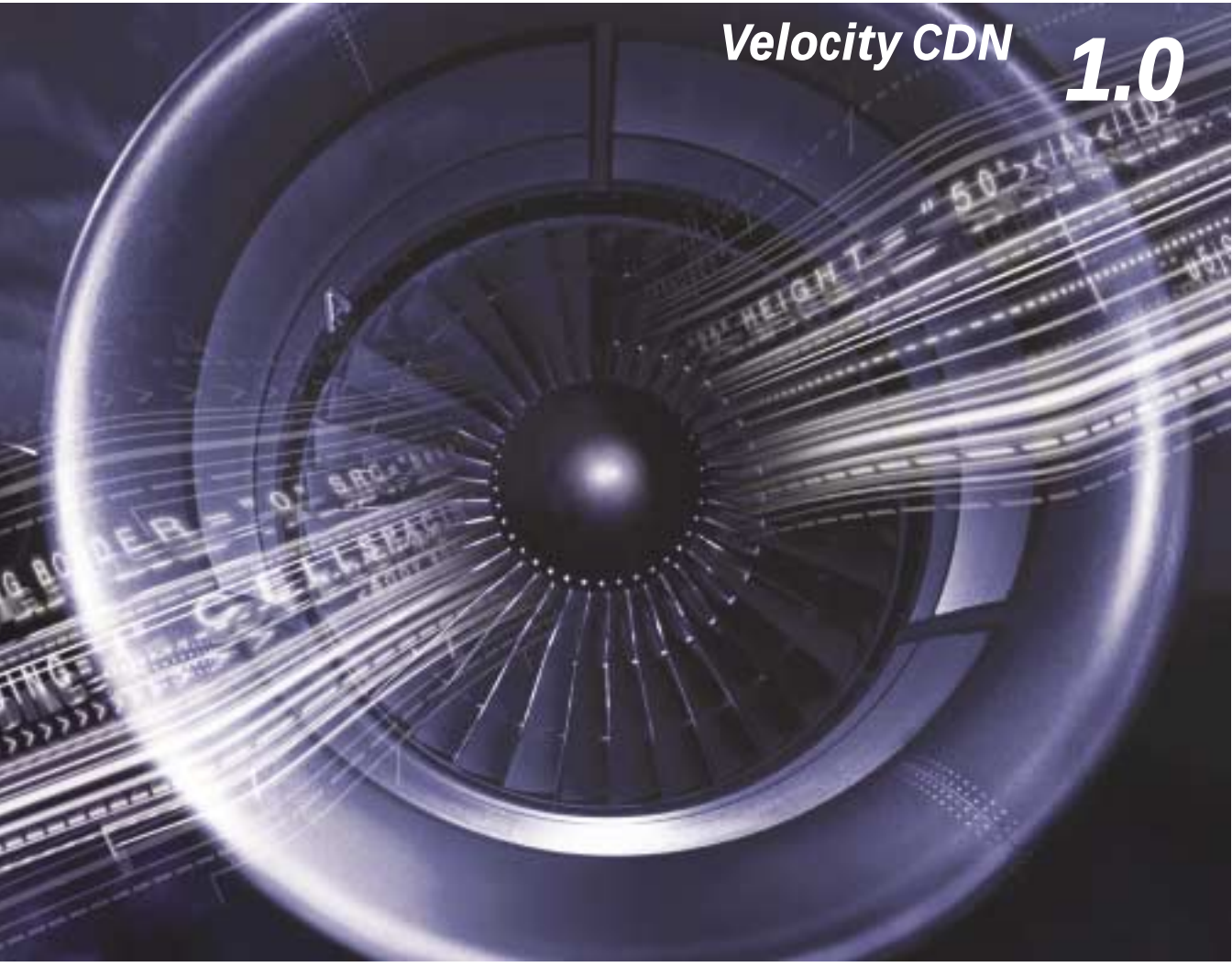
Volera™
Make the Net Fly



Aggregating and Reporting Statistics with Content Accountant

Legal Notices

Volera, Inc.
2211 North First Street
San Jose, CA 95131-2021
U.S.A.

www.volera.com

Aggregating and Reporting Statistics with Content Accountant
August 2001

Online Documentation: To access the online documentation for this and other Volera products, and to get updates, see www.volera.com.

Volera Trademarks

Volera is a trademark of Volera, Inc. in the United States and other countries.

Third-Party Trademarks

All third-party trademarks are the property of their respective owners.

Contents

About this Guide 7

1 Content Accountant Overview 9

CA Components 10

2 Configuring Content Accountant 11

System Requirements 11
How CA Server Scales Its Available Resources 11
Aggregation Happens at Two Points 12
Changing the Content Accountant Configuration 13
Configuring Content Accountant to Report Statistics 13
Monitoring Content Accountant Performance 13
Velocity CDN 1.0 Content Accountant Reports in One of Three Formats 13

3 Integrating Content Accountant with Third-Party Billing Systems 15

4 Integrating with Third-Party Reporting/Analytical Systems 17

About this Guide

This guide contains information regarding Velocity CDN™ 1.0 Content Accountant.

To	See
Learn how Content Accountant works	Chapter 1, “Content Accountant Overview,” on page 9
Configure Content Accountant	Chapter 2, “Configuring Content Accountant,” on page 11
Integrate Content Accountant with third-party billing systems	Chapter 3, “Integrating Content Accountant with Third-Party Billing Systems,” on page 15
Integrate Content Accountant with third-party reporting/analytical systems	Chapter 4, “Integrating with Third-Party Reporting/ Analytical Systems,” on page 17

1 Content Accountant Overview

Content Accountant lets you track content delivery and document a variety of CDN statistics. Data is filtered so that the information that is important to you is collected.

Because Content Accountant can be installed independently from System Controller, a CDN can have as many instances of the Content Accountant Server as are required to aggregate and report content usage statistics.

The results of Content Accountant reports can be integrated with a variety of billing and reporting systems, including Apogee and Portal to provide truly meaningful reporting. And for data analysis, you can also integrate Content Accountant with WebTrends, Analog, Aria, and Accrue.

Content Accountant lets you determine which users are accessing what content in real time. This lets you decide what type of content works best for users, so you can plan to create and deploy the optimal content type.

Content Accountant (CA) collects the following four billing-related statistics from one or more remote Excelerator systems on a per-object basis:

- Total hits
- Total misses
- Total bytes served from cache
- Total bytes filled from origin

Content Agent then organizes, stores, and exports these statistics to as many report-generating portals as you specify in a single data transmission.

Instructions for installing Content Accountant are in *Getting Started with Content Accountant*.

CA Components

For a high-level overview of Content Accountant, see Content Accountant Overview in *Planning Your Content Distribution Network (CDN)*.

CA consists of two components:

- **CA Agent:** A statistics collection agent that runs on Excelerator 2.1 caching servers.

  CA Agent is fully integrated with Excelerator 2.1 and load automatically during system startup. It remains idle until it receives one or more jobs from Content Controller. CA Agent requires an active Content Controller License.

  This receives administrator-created jobs from Content Controller. Jobs consist of one or more content collection definitions and a set of options.

  It collects Excelerator statistics for these content sets based on the collection frequency and immediately transmits the results to the CA Server.

- **CA Server:** An Apache module/Java application that runs on the VCDN management server.

  CA Server must be installed and configured using the CA Server product CD. No license is required.

  This aggregates data from multiple Excelerator devices and transfers it to one or more third-party billing/trending systems, such as Apogee, Portal, WebTrends, Analog, Aria, and Accrue.

2 Configuring Content Accountant

This chapter contains information regarding configuration of Content Accountant.

To	See
Understand the requirements for running Content Accountant	“System Requirements” on page 11
Understand how Content Accountant automatically scales it resource consumption to match its workload	“How CA Server Scales Its Available Resources” on page 11
Change Content Accountant’s initial configuration settings	“Changing the Content Accountant Configuration” on page 13
Monitor Content Accountant performance	“Monitoring Content Accountant Performance” on page 13

System Requirements

Content Accountant requires that Content Controller is functioning properly. For more information on setting up Content Controller, see Creating Content Collections and Jobs in *Managing Content with Content Controller*.

How CA Server Scales Its Available Resources

CA Server must be able to service HTTP posts from up to 1,000 Excelerator caching servers that, in turn, might be monitoring numerous collections and objects. CA Server must also be able to transmit the statistics it collects frequently.

To meet these requirements, CA Server will automatically scale its available hardware resources to one of the scenarios outlined in the following Table:

Table 1 CA Server Scaling Scenarios

Scenario	Caching Servers	Collections	Objects	Collection/Transmission Frequency
Low	Up to 1,000	10	100	24 hours
Medium	Up to 1,000	100	1,000	1 hour
High	Up to 1,000	1,000	10,000	30 minutes

Aggregation Happens at Two Points

Each cache device aggregates its statistics before sending them to Content Accountant Server. Then Content Accountant Server also aggregates the statistics from all the devices before sending them to the FTP server.

Statistics are sent to the FTP server in one of three formats:

- SQUID

  This format stores a log entry for each request.

- Extended W3C

  This format stores a log entry for each request.

- Summarized W3C

  This format reports aggregated statistics

SQUID and Extended W3C are useful for CDNs that need to track and report on single objects in small collections. For example, if you are tracking large media files and need to know exact information regarding each object usage, SQUID or W3C would provide that information. If large numbers of object requests are logged, the data collected becomes very large very quickly and can consume large amounts of network bandwidth for transmission. For example, 5,000 requests for a single object result in 5,000 entries in the logged data.

Summarized W3C is useful for CDNs that have large sites with many objects. Because content usage data is summarized, bandwidth required for data transfers is less by orders of magnitude. For example, 5,000 requests for a single object result in one log entry.

Changing the Content Accountant Configuration

Configuration changes to Content Accountant are made by editing the /etc/rc.d/init.d/httpd file. This might be required if the IP configuration of the system changes or additional third-party aggregating and reporting services are installed.

Configuring Content Accountant to Report Statistics

When you create a job for managing content collections, if you have Content Accountant installed on your network, you can specify the Content Accountant server to which cache devices should report usage statistics for objects in the collection and the destination (FTP) server to which Content Accountant server will send the aggregated statistics.

Monitoring Content Accountant Performance

You can monitor Content Accountant Agent by visually checking the management tool's management views or by manually verifying sample data.

You can verify Content Accountant Server health by manually validating the server's output.

Velocity CDN 1.0 Content Accountant Reports in One of Three Formats

3 Integrating Content Accountant with Third-Party Billing Systems

This chapter will contain pointers for configuring CA to work with third-party systems, formats, filtering, scheduling, etc.

4 Integrating with Third-Party Reporting/Analytical Systems

This chapter will contain information regarding integrating Content Accountant with third-party reporting/analytical systems.