EMC Celerra Version 5.6 Technical Primer: Celerra Event Publishing Agent (CEPA) for Content Quota Management

Technology Concepts and Business Considerations

Abstract

This white paper introduces the Celerra® Event Publishing Agent. EMC has enabled seamless integration between Celerra and products from leading content/quota management partners, extending overall management and administration capabilities.

June 2009
Executive summary

Maintaining efficient usage of storage resources is a daily challenge for a data center. To assist in this effort, EMC has introduced Celerra® Event Publishing Agent (CEPA) in Celerra version 5.6. This new capability enables seamless integration between Celerra and industry-leading content/quota management software applications (CQM applications). Together, this combination provides an enhanced set of functionality that extends the management capability of Celerra and delivers improved utilization.

Business problem

Maximizing storage usage is a top priority for CIOs and storage administrators. Inefficient allocation or use of storage resources can prove to be rather costly. For example, poor utilization can lead to the procurement of new storage systems that would otherwise be unnecessary. Additionally, expanding storage infrastructures increases the demand on IT administration teams as well as on network and backup resources. The cost of infrastructure management is significant; it is commonly reported that for every dollar spent on storage it costs five to seven dollars to manage it.

Addressing storage usage by unstructured data (for example, user files) is generally a primary goal because unstructured data generally accounts for 85 percent of all data stored. These files normally include valid business data created by applications such as MS Office (for example, Word, Excel, PowerPoint) or Adobe Acrobat (for example, PDF). However, it also includes non-vital, obsolete, or even policy-violating user data (for example, music files, photos), which reports claim can grow to 30 percent of all data stored when left unmanaged.

Effectively addressing these issues can have a powerful impact on the efficiency and performance of most businesses, but taking action is difficult without the proper tools. Changing user behavior and manually gathering the necessary information to take action is laborious and time consuming. However, you can surmount these challenges using the Celerra Event Publishing Agent and a leading CQM application. For the first time, storage administrators now have the advanced and automated tools they need to successfully set and enforce usage policies in their environments.

Technical problem

Although the business benefits of storage resource management (SRM) are widely acknowledged, the task of taking action using SRM data can be challenging. The process of identifying files that are in violation of corporate policy, unnecessary, or undesired can be overwhelming. Even if content can be identified it still may be difficult to approach or modify the behavior of offending users without the proper tools.

The challenge then becomes how to address this issue in an efficient manner to allow an IT group to viably add it to an already long list of existing activities.

Celerra Event Publishing Agent in Celerra version 5.6, along with a qualified CQM application, is meant to address this challenge, placing the appropriate tools in the hands of the IT staff.

Feature introduction

Celerra Event Publishing Agent in Celerra version 5.6 is a new feature that enables integration with leading CQM applications. This combination enhances existing Celerra functionality while providing customers the freedom to choose the CQM application that best suits their specific needs for quota and content management.

---

1  NORTHERN Parklife, Inc., Northern Storage Suite V8, NTP Software, NTP Software QFS
What's new
Celerra Event Publishing Agent is a new feature included in Celerra version 5.6. It is comprised of a host framework and agent that have been validated to support several leading CQM application vendors.

This integration has advantages over similar offerings available today including:
- **Faster response on CIFS**: Optimized component communication delivers new efficiencies.
- **High availability**: Distributed architecture is fault tolerant.
- **Fail-safe operation**: Solution design ensures an uninterrupted client user experience.
- **CAVA compatibility**: The integrated design provides seamless integration between CEPA and the popular Celerra AntiVirus Agent (CAVA).
- **Native 64-bit support in Windows**: Support for Windows servers such as Vista and Windows Server 2008 was introduced in version 5.6.45.x.

When deployed end users experience a new level of operational control, including:
- **Comprehensive quota management** (users, groups, hard limits, soft limits with user notification, threshold event triggers)
- **Storage usage dashboards** (current state snapshots, warnings on current and future projected issues, recommended actions)
- **In-depth administrative reporting** (space allocation, growth projections, drill downs, file attributes, scheduling)
- **Proactive user reporting and chargeback** (auto-forward usage and compliance reports to users)
- **Policy enforcement** (user adherence to business rules, for example, limiting MP3s)
- **Event triggering** (launch actions based upon observed activity or threshold breach)

Introduction
This white paper provides an overview of EMC Celerra Event Publishing Agent and helps the reader develop a functional understanding of the product and technology. Information includes:

- Business case for CEPA
- Deployment workflows
- Architecture of CEPA
- Limitations

This white paper was formerly titled *EMC Celerra Version 5.6 Technical Primer: Celerra Event Publishing Agent (CEPA) – Technology Concepts and Business Considerations*.

Audience
The intended audience is customers, including IT planners, storage architects and administrators, and any others involved in evaluating, acquiring, managing, operating, or designing an EMC networked storage environment.

Terminology
**Celerra Event Enabler Framework (CEE)**: Windows host agent software that contains both the Celerra Event Publishing Agent (CEPA) and Celerra AntiVirus Agent (CAVA) features. CEE must be installed on all CEPA Servers.
**Celerra Event Publishing Agent (CEPA):** Functionality contained within CEE that enables Celerra integration with select CQM applications.

**Content/Quota Management Software (CQM Application):** Third-party content/quota management applications that may individually be used with the Celerra Event Publishing Agent to deliver advanced management functionality.

**CEPA Server:** Windows host server(s) on which CEPA and CEE have been installed. All CEPA Servers must be members of the same Windows domain.

**CQM Host:** A Windows host server on which the CQM application is installed. A CQM Host may also be a CEPA Server (that is, the CQM application may be co-resident with CEPA). CQM Hosts must be part of the same Windows domain as the CEPA server(s) assigned to communicate to it.

**CEPA Pool:** A list (that is, configuration file) maintained within each Celerra Data Mover that identifies the CEPA Server(s) that are hosting the CEPA application. Each Celerra Data Mover uses this list in its round-robin schedule processing.

**CEPA Pool Member:** A CEPA Server identified in a CEPA Pool list.

**Celerra Data Mover (Data Mover):** A Celerra Network Server cabinet component running its own operating system that retrieves files from a storage device and makes them available to a network client. A Celerra Network Server may contain several Data Movers. Any Data Mover configured for use with CEPA maintains its own independent CEPA Pool list.

**CIFS (Common Internet File Service):** A file-sharing protocol based on the Microsoft Server Message Block (SMB) application-level network protocol. The CIFS protocol is primarily used for file sharing by Windows platforms.

**Celerra Command Line Interface (CLI):** A text-based interface used for executing Celerra system commands.

**User Client Request:** A user-initiated request attempts an action such as the following: create, open, close, close after modification, rename, or delete.

**Round-robin Scheduling:** A sequential scheduling algorithm that assigns equal priority to tasks across available resources.
**Detailed overview**

Implementation of Celerra with the Celerra Event Publishing Agent and a qualified CQM application greatly enhances a system administrator’s ability to manage and control usage of critical storage resources.

**Sample deployment workflows**

Figure 1, Figure 2, and Figure 3 show some sample deployment workflows.

---

**Figure 1.** A single Data Mover, single CEPA Server configuration with a co-resident CQM application

---

**Figure 2.** A single Data Mover, single CEPA Server configuration with an external CQM Host
Figure 3. A single Data Mover, multi-CEPA Server configuration with multiple CQM Hosts

CEPA has the versatility of being configured in many ways; however, the steps for CEPA/CQM are always the same. Steps 1 through 6 correlate to each figure above.

1. **User Client Request** - A User Client Request is initiated by a user.

2. **CEPA Request** - Upon receipt, the Data Mover processes the request and, if required, forwards the request to one of the CEPA Servers (that is, CEPA Pool Members) identified on its CEPA Pool list.

3. **CEPA Forwarding** - The identified CEPA Server (that is, CEPA Pool Member) then forwards the request to the one or more CQM applications registered for that request.

4. **CQM Application Response** – The CQM application processes the request and returns the policy response to the CEPA Server that requested it.

5. **CEPA Response** - Upon receipt of the CQM application response the CEPA Server (that is, CEPA Pool Member) forwards the response to the originating Data Mover. This response ensures that the Celerra Data Mover takes an action consistent with previously defined CQM application parameters.

6. **Celerra Response to User Client Request** – The Celerra Data Mover returns a response communicating the outcome of the initial User Client Request (that is, action #1).

   a. **Success:**
      i. **Client User Write Request did not** require CEPA/CQM review and was executed successfully by Celerra.
      ii. **Client User Write Request did** require CEPA/CQM review, was granted approval by CEPA/CQM, and was executed successfully by Celerra.
   b. **Failure:**
      i. **Client User Write Request did not** require CEPA/CQM review but was not executed successfully by Celerra for an unrelated reason.
      ii. **Client User Write Request did** require CEPA/CQM review, was not granted approval by CEPA/CQM, and was not executed successfully by Celerra.
**Architecture**

A typical CEPA solution (as shown in Figure 2 on page 7) includes the following elements. Other aspects of the solutions such as scalability and performance are discussed in the sections that follow:

- **Solution hardware components**
  - Celerra Network Server running Celerra version 5.6
  - CEPA Server – host for CEE (CEPA)
    - Multiple CEPA Servers may be deployed for redundancy and load balancing.
  - CQM Host
  - Windows client
  - Windows domain server
  - Network connectivity between all components
- **Solution software components:**
  - Celerra Event Enabler Framework (CEE), containing CEPA
  - CQM application (Available for order via EMC Product Catalog)

![Figure 4. Architecture of a single CEPA Server with a co-resident CQM application](image)

**Communication protocol**

- The physical connectivity between the various components in this solution is Ethernet.

Note: All CEPA solution components (CEPA Server(s), Windows domain controller, CQM Host) must participate in the same or trusted Windows domain environment. This is required in order to authenticate the individual components.

- Remote procedure calls (RPC) are used to transport messages over the physical layer between solution components. The RPC network port is 111 and must be opened in the network firewall.
- Local remote procedure calls (LRPC) are additionally used in the case where the CQM application is co-resident on a CEPA Server.
- To support native 32-bit and 64-bit Windows, Celerra to CEPA communications now uses MS-RPC.
- The new version of CEPA is backward-compatible.

**Software architectural overview**

- Celerra Data Mover processes CIFS requests in accordance with policies set within the CQM application to determine which requests require policy/quota management processing.
  - User Client Requests requiring processing are forwarded by the Data Mover to one of the CEPA Servers identified on the CEPA Pool list. (If more than one CEPA Server is deployed Celerra will alternate (for example, round-robin) its target CEPA Server to load balance the workload.)
CEPA is activated on a per-Celerra Data Mover basis. On activated CEPA Data Movers, CEPA is enabled for every mounted file system, except root file systems by default, unless explicitly excluded by an administrator.

- CEE is the EMC software that must be installed on each CEPA Server. It ensures proper communication between Celerra and the CEPA solution components.
- The CQM application, which can be co-resident on one of the CEPA Servers or another host within the Windows domain environment, processes the requests based on management policies defined in the CQM application.
- Messaging Service is the communication protocol used to communicate events to and from CEPA. It is an application protocol that is transmitted over the RPC transport protocol.
- Heartbeat is a mechanism that Celerra uses to monitor each CEPA solution component:
  - CEPA Server - The status of each CEPA Server is monitored for viable participation in the CEPA Pool that is used by Celerra to load balance requests.
  - CQM Host – The status of the CQM host/CQM application is monitored for viable event processing capability.
  - If connectivity is lost, preventing CQM application processing of CEPA Requests, Celerra will assume a state predefined by the administrator:
    - DENY - The Celerra Data Mover blocks all User Client Requests until connectivity is restored
    - IGNORE - The Celerra Data Mover allows all User Client Requests until connectivity is restored. When connectivity is restored an “administrative event” is sent to the CQM application documenting the number of User Client Requests allowed during the affected period.

Scalability
- The CEPA solution scales by adding CEPA Servers that incrementally improve processing speed and response times to client applications. Each additional CEPA Server must be manually added to CEPA Pool lists so that each Data Mover identifies the server as a valid CEPA Pool Member.
- A round-robin algorithm is used to assign processing requests to CEPA Pool Members of the CEPA Pool list. This approach distributes the processing workload and improves response times to client applications.

Security
- The CEPA solution has been designed to provide the highest possible level of security. It uses remote procedure calls (RPC) for solution component communication and Active Directory authentication and packet level integrity.
  - Local administration – A local administrator can be defined on the Windows Domain Controller to manage CEPA on the CEPA Server(s). The administrator lists the endpoints (that is, IP addresses of the CQM Hosts) that CEPA will actively communicate with using RPC. Each CQM application will register with the CEPA Server(s), providing policies that CEPA will monitor and forward for processing. Using the same communication, the CQM application delivers a message back with its response to the triggered event.
  - CEPA solution component authentication – For proper authentication CEPA Servers must reside in the same Windows domain as the CQM Host to which they communicate. This does not preclude deployments across multiple Windows domains. In such cases any CEPA Server(s) communicating to a given CQM Host would be required to reside in the same Windows domain of that specific CQM Host.
  - Denial-of-service attacks can be avoided by local administrators by using “option=ignore/denied” to regulate CIFS traffic when a CEPA Server might be unresponsive (see the statement on “Heartbeat” in the “Software architectural overview” section).
Performance

- CEPA is designed to deliver performance improvements over a non-integrated, standalone CQM application deployment.
  - Communication is more efficient. CEPA intelligently packages event trigger and event context data into a single communication, minimizing dialogue with the CQM application.
  - Use of ACLS is not required. CEPA does not require the CQM application to rely on ACLS for data access.
- Quantifying expected performance is difficult but will be a function of CQM application processing time of a given request and the quantity of CEPA Pool Members used in the solution.

Licensing

Licensing a CEPA solution requires the following components:

- CQM Application License – available via the EMC Partner Select Program.
- Celerra Event Enabler infrastructure (model number NSxx-CEE-L) – required for every CEPA Server used in the deployment. All required CEPA software elements are included within the Celerra Event Enabler infrastructure.

Limitations

Deployment

- Only Windows environments are supported. At this time CQM applications are limited to Windows (that is, CIFS) environments.
- The configuration of the CEPA solution is performed only through the Celerra CLI.
- A native 64-bit installation of CEPA will not work on a 32-bit Windows server.
- A native 32-bit installation of CEPA will not work on a 64-bit Windows server.

Network protocol

- Only Windows (that is, CIFS) client requests can be processed at this time.

Compatibility with older releases

The CEPA solution is valid only on Celerra version 5.6 and later software revisions.
Conclusion

EMC Celerra customers can manage their content and quota environment from one CQM application. Together CEPA and CQM deliver faster response on CIFS, in-depth administrative reporting, and realtime policy enforcement.

References

The following documents can be found on Powerlink®, EMC’s password-protected, customer- and partner-only extranet:

- **Using Celerra Event Enabler**: This document describes how to install the EMC Celerra Event Enabler framework, which provides a working environment for Celerra AntiVirus Agent (CAVA) and Celerra Event Publishing Agent (CEPA).

- **Using Celerra Event Publishing Agent (CEPA)**: This document describes how to use the Celerra Event Publishing Agent (CEPA), which is a mechanism whereby third-party applications can register to receive event notification and context from Celerra. CEPA delivers to the consuming application both event notification and associated context in one message.

- Man pages/command descriptions (reflected in the *EMC Celerra Network Server Command Reference Manual*)

- **Getting Started with Celerra Event Enabler**: This document provides an overview of the CEE software kit and documentation set.