



Reporting

Cloud Insights

NetApp
October 26, 2021

This PDF was generated from https://docs.netapp.com/us-en/cloudinsights/reporting_overview.html on October 26, 2021. Always check docs.netapp.com for the latest.

Table of Contents

Reporting	1
Cloud Insights Reporting Overview	1
Cloud Insights Reporting User Roles	2
Predefined Reports Made Easy	3
Storage Manager Dashboard	7
Creating a Report (Example)	9
Managing Reports	11
Creating Custom Reports	13
How historical data is retained for Reporting	20
Cloud Insights Data Warehouse Schema Diagrams	21
Cloud Insights Schemas for Reporting	62

Reporting

Cloud Insights Reporting Overview

Cloud Insights reporting is a business intelligence tool that enables you to view pre-defined reports or create custom reports.



The Reporting feature is available in Cloud Insights [Premium Edition](#).

With Cloud Insights reporting you can perform the following tasks:

- Run a pre-defined report
- Create a custom report
- Customize a report's format and delivery method
- Schedule reports to run automatically
- Email reports
- Use colors to represent thresholds on data

Cloud Insights Reporting can generate custom reports for areas like chargeback, consumption analysis, and forecasting, and can help answer questions such as the following:

- What inventory do I have?
- Where is my inventory?
- Who is using our assets?
- What is the chargeback for allocated storage for a business unit?
- How long until I need to acquire additional storage capacity?
- Are business units aligned along the proper storage tiers?
- How is storage allocation changing over a month, quarter, or year?

Accessing Cloud Insights Reporting

You can access Cloud Insights Reporting by clicking the **Reports** link in the menu.

You will be taken to the Reporting interface. Cloud Insights uses IBM Cognos Analytics for its reporting engine. Log in using your Reporting credentials.

What is ETL?

When working with Reporting, you will hear the terms "Data Warehouse" and "ETL". ETL stands for "Extract, Transform, and Load". The ETL process retrieves data collected in Cloud Insights, and transforms the data into a format for use in Reporting. "Data Warehouse" refers to the collected data available for Reporting.

The ETL process includes these individual processes:

- **Extract:** Takes data from Cloud Insights.
- **Transform:** Applies business logic rules or functions to the data as it is extracted from Cloud Insights.

- **Load:** Saves the transformed data into the data warehouse for use in Reporting.

Cloud Insights Reporting User Roles

If you have Cloud Insights Premium Edition with Reporting, every Cloud Insights user in your environment also has a Single Sign-On (SSO) login to the Reporting application (i.e. Cognos). Simply click the **Reports** link in the menu and you will automatically be logged in to Reporting.

Your user role in Cloud Insights determines your Reporting user role:

Cloud Insights Role	Reporting Role	Reporting Permissions
Guest	Consumer	Can view, schedule, and run reports and set personal preferences such as those for languages and time zones. Consumers cannot create reports or perform administrative tasks.
User	Author	Can perform all Consumer functions as well as create and manage reports and dashboards.
Administrator	Administrator	Can perform all Author functions as well as all administrative tasks such as configuration of reports and the shutdown and restart of reporting tasks.

The following table shows the functions available to each Reporting role.

Feature	Consumer	Author	Administrator
View reports in the Team Content tab	Yes	Yes	Yes
Run reports	Yes	Yes	Yes
Schedule reports	Yes	Yes	Yes
Upload external files	No	Yes	Yes
Create Jobs	No	Yes	Yes
Create stories	No	Yes	Yes
Create reports	No	Yes	Yes
Create Packages and Data Modules	No	Yes	Yes
Perform administrative tasks	No	No	Yes

Setting Reporting (Cognos) email preferences



If you change your user email preferences within Cloud Insights Reporting (i.e. the Cognos application), those preferences are active *only for the current session*. Logging out of Cognos and back in again will reset your email preferences.

Important note for existing customers

If you are new to Cloud Insights with Reporting, welcome! There is nothing more you need to do to begin enjoying Reporting.

If you are a current Premium Edition customer, SSO is not automatically enabled for your environment. When you enable SSO, the administrator user for the reporting portal (Cognos) ceases to exist. This means that any reports that are in the *My Content* folder are removed and must be reinstalled or re-created in *Team Content*. Additionally, scheduled reports will need to be configured once SSO is enabled.

What steps should I take to prepare my existing environment for enabling SSO?

To ensure your reports are retained, migrate all reports from *My Content* to *Team Content* using the following steps. You must do this prior to enabling SSO in your environment:

1. Create a new folder in *Team Content*
 - a. If multiple users have been created, please create a separate folder for each user to avoid overwriting reports with duplicate names
2. Navigate to *My Content*
3. Select all of the reports you wish to retain.
4. In the upper right corner of the menu, select "Copy or move"
5. Navigate to the newly created folder in *Team Content*
6. Paste the reports to the newly created folder using the "Copy to" or "Move to" buttons
7. Once SSO is enabled for Cognos, log into Cloud Insights with the email address used to create your account.
8. Navigate to the *Team Content* folder within Cognos, and Copy or Move the previously saved reports back to *My Content*.

Predefined Reports Made Easy

Cloud Insights Reporting includes predefined reports that address a number of common reporting requirements, providing critical insight that stakeholders need to make informed decisions about their storage infrastructure.



The Reporting feature is available in Cloud Insights [Premium Edition](#).

You can generate pre-defined reports from the Cloud Insights Reporting Portal, email them to other users, and even modify them. Several reports enable you to filter by device, business entity, or tier. The reporting tools use IBM Cognos as a foundation and give you many data presentation options.

The pre-defined reports show your inventory, storage capacity, chargeback, performance, storage efficiency, and cloud cost data. You can modify these pre-defined reports and save your modifications.

You can generate reports in various formats, including HTML, PDF, CSV, XML, and Excel.

Cloud Insights accommodates multiple tenancy in reporting by enabling you to associate users with business units. With this feature, administrators can separate data or reports according to the attributes of a user or his/her affiliation.

Navigating to Pre-defined Reports

When you open the Reporting Portal, the *Team Content* folder is the starting point for you to select the type of information that you require in the Cloud Insights reports.

1. In the left navigation pane, click **Team Content** and select the information category that you want to use.
2. Click **Reports** to access the pre-defined reports.

The screenshot shows the IBM Cognos Analytics Reporting Portal. The left sidebar has links for Home, Search, My content, Team content (selected), and Recent. The main content area shows a list under 'Team content' with three items: Packages (3/21/2019 6:47 PM), Reports (3/21/2019 10:44 PM), and Storage Manager Dashboard (4/16/2019 7:09 PM).

Using predefined reports to answer common questions

The following predefined reports are available in **Team content > Reports**.

Application Service Level Capacity and Performance

The Application Service Level Capacity and Performance report provides a high level overview of your applications. You can use this information for capacity planning or for a migration plan.

Chargeback

The Chargeback report provides storage capacity chargeback and accountability information by hosts, application, and business entities, and includes both current and historical data.

To prevent double counting do not include ESX servers, only monitor the VMs.

Data Sources

The Data Sources report shows all the data sources that are installed on your site, the status of the data source (success/failure), and status messages. The report provides information about where to start

troubleshooting data sources. Failed data sources impact the accuracy of reporting and the general usability of the product.

ESX vs VM Performance

The ESX vs VM Performance report provides a comparison of ESX servers and VMs, showing average and peak IOPs, throughput, and latency and utilizations for ESX servers and VMs. To prevent double counting, exclude the ESX servers; only include the VMs.

An updated version of this report is available at the NetApp Storage Automation Store.

Fabric Summary

The Fabric Summary report identifies switches and switch information, including port counts, firmware versions, and license status. The report does not include NPV switch ports.

Host HBAs

The Host HBAs report provides an overview of the hosts in the environment and provides the vendor, model, and firmware version of HBAs, and the firmware level of the switches to which they are connected. This report can be used to analyze firmware compatibility when planning a firmware upgrade for a switch or an HBA.

Host Service Level Capacity and Performance

The Host Service Level Capacity and Performance report provides an overview of storage utilization by host for block only applications.

Host Summary

The Host Summary report provides an overview of storage utilization by each selected host with information for Fibre Channel and iSCSI hosts. The report enables you to compare ports and paths, the Fibre Channel and iSCSI capacity, and violation counts.

License Details

The License Details report shows the entitled quantity of resources you are licensed for across all sites with active licenses. The report also shows a summation of actual quantity across all the sites with active licenses. The summation may include overlaps of storage arrays managed by multiple servers.

Mapped but not Masked Volumes

The Mapped but not Masked Volumes report lists the volumes whose logical unit number (LUN) has been mapped for use by a particular host, but is not masked to that host. In some cases these could be decommissioned LUNs that have been unmasked. Unmasked volumes can be accessed by any host, making them vulnerable to data corruption.

NetApp Capacity and Performance

The NetApp Capacity and Performance report provides global data for allocated, utilized, and committed capacity with trending and performance data for NetApp capacity.

Scorecard

The Scorecard report provides a summary and general status of all assets acquired by Cloud Insights. Status is indicated with green, yellow, and red flags:

- Green indicates normal condition
- Yellow indicates a potential issue in the environment
- Red indicates an issue that requires attention

All of the fields in the report are described in the Data Dictionary provided with the report.

Storage Summary

The Storage Summary report provides a global summary of used and unused capacity data for raw, allocated, storage pools, and volumes. This report provides an overview of all of the storage discovered.

VM Capacity and Performance

Describes the virtual machine (VM) environment and its capacity usage. VM tools must be enabled to view some data, such as when VMs were powered down.

VM Paths

The VM Paths report provides data store capacity data and performance metrics for which virtual machine is running on which host, which hosts are accessing which shared volumes, what the active access path is, and what comprises capacity allocation and usage.

HDS Capacity by Thin Pool

The HDS Capacity by Thin Pool report shows the amount of usable capacity on a storage pool that is thin provisioned.

NetApp Capacity by Aggregate

The NetApp Capacity by Aggregate report shows raw total, total, used, available, and committed space of aggregates.

Symmetrix Capacity by Thick Array

The Symmetrix Capacity by Thick Array report shows raw capacity, useable capacity, free capacity, mapped, masked, and total free capacity.

Symmetrix Capacity by Thin Pool

The Symmetrix Capacity by Thin Pool report shows raw capacity, useable capacity, used capacity, free capacity, used percentage, subscribed capacity, and subscription rate.

XIV Capacity by Array

The XIV Capacity by Array report shows used and unused capacity for the array.

XIV Capacity by Pool

The XIV Capacity by Pool report shows used and unused capacity for storage pools.

Storage Manager Dashboard

The Storage Manager Dashboard provides you with a centralized visualization that enables you to compare and contrast resource usage over time against the acceptable ranges and previous days of activity. Showing only the key performance metrics for your storage services, you can make decisions about how to maintain your data centers.



The Reporting feature is available in Cloud Insights Premium Edition.

The dashboard comprises seven components that contain contextual information on certain aspects of your storage environment. You can drill down on the aspects of your storage services to perform an in-depth analysis of a section that interests you most.

Summary



This component shows the used versus usable storage capacity, total switch ports versus the number of switch ports connected, and total connected switch port utilization versus the total bandwidth, and how each of these trend over time. You can view the actual utilization compared against the low, mid, and high ranges, which enables you to compare and contrast usage between projections and your desired actuals, based on a target. For capacity and switch ports, you can configure this target. The forecast is based on an extrapolation of the current growth rate and the date you set. When the forecasted used capacity, which is based on future usage projection date, exceeds the target, an alert (solid red circle) appears next to Capacity.

Storage Tiers Capacity

This component shows the tier capacity used versus the capacity allocated to the tier, which indicates how the used capacity increases or decreases over a 12-month period and how many months are remaining to full capacity. Capacity usage is shown with values provided for actual usage, the usage forecast, and a target for capacity, which you can configure. When the forecasted used capacity, which is based on future usage projection date, exceeds the target capacity, an alert (solid red circle) appears next to a tier.

You can click any tier to display the Storage Pools Capacity and Performance Details report, which shows free versus used capacities, number of days to full, and performance (IOPS and Response Time) details for all the pools in the selected tier. You can also click any storage or storage pool name in this report to display the asset page summarizing the current state of that resource.

Daily Storage Traffic

This component shows how the environment is performing, if there is any large growth, changes, or potential issues compared to the previous six months. It also shows the average traffic versus the traffic for the previous seven days, and for the previous day. You can visualize any abnormalities in the way the infrastructure is performing because it provides information that highlights both cyclical (previous seven days) and seasonal variations (previous six months).

You can click the title (Daily Storage Traffic) to display the Storage Traffic Details report, which shows the heat map of the hourly storage traffic for the previous day for each storage system. Click any storage name in this report to display the asset page summarizing the current state of that resource.

Data Centers Time to Full

This component shows all the data centers versus all of the tiers and how much capacity remains in each data center for each tier of storage based on forecasted growth rates. Tier capacity level is shown in blue; the darker the color, the lesser time the tier at the location has left before it is full.

You can click a section of a tier to display the Storage Pools Days to Full Details report, which shows total capacity, free capacity, and number of days to full for all the pools in the selected tier and the data center. Click any storage or storage pool name in this report to display the asset page summarizing the current state of that resource.

Top 10 Applications

This component shows the top 10 applications based on the used capacity. Regardless of how the tier organizes the data, this area displays the current used capacity and share of the infrastructure. You can visualize the range of user experience for the previous seven days to see if consumers experience acceptable (or, more importantly, unacceptable) response times.

This area also shows trending, which indicates if the applications meet their performance service level objectives (SLO). You can view the previous week's minimum response time, the first quartile, the third quartile, and the maximum response time, with a median shown against an acceptable SLO, which you can configure. When the median response time for any application is out of the acceptable SLO range, an alert (solid red circle) appears next to the application. You can click an application to display the asset page summarizing the current state of that resource.

Storage Tiers Daily Performance

This component shows a summary of the tier's performance for response time and IOPS for the previous seven days. This performance is compared against a SLO, which you can configure, enabling you to see if there is opportunity to consolidate tiers, realign workloads delivered from those tiers, or identify issues with

particular tiers. When median response time or median IOPS is out of the acceptable SLO range, an alert (solid red circle) appears next to a tier.

You can click a tier name to display the Storage Pools Capacity and Performance Details report, which shows free versus used capacities, number of days to full, and performance (IOPS and response time) details for all the pools in the selected tier. Click any storage or storage pool in this report to display the asset page summarizing the current state of that resource.

Orphaned Capacity

This component shows the total orphaned capacity and orphaned capacity by tier, comparing it against acceptable ranges for total usable capacity and showing the actual capacity that is orphaned. Orphaned capacity is defined by configuration and by performance. Storage orphaned by configuration describes a situation in which there is storage allocated to a host. However, the configuration has not been performed properly and the host cannot access the storage. Orphaned by performance is when the storage is correctly configured to be accessed by a host. However, there has been no storage traffic.

The horizontal stacked bar shows the acceptable ranges. The darker the gray, the more unacceptable the situation is. The actual situation is shown with the narrow bronze bar that shows the actual capacity that is orphaned.

You can click a tier to display the Orphaned Storage Details report, which shows all the volumes identified as orphaned by configuration and performance for the selected tier. Click any storage, storage pool, or volume in this report to display the asset page summarizing the current state of that resource.

Creating a Report (Example)

Use the steps in this example to generate a simple report on physical capacity of storage and storage pools in a number of data centers.

Steps

1. In the toolbar, click **[+]**
2. Click **Report**
3. Click **Templates > Blank**
4. Click **Themes > Cool Blue > OK**

The Source and Data tabs is displayed

5. Click **Source > [+]**
6. In the **Open file** dialog, click **Team content > Packages**

A list of available packages is displayed.

7. Click **Storage and Storage Pool Capacity > Open**
8. Click **[+]**

The available styles for your report are displayed.

9. Click **List**

Add appropriate names for List and Query

10. Click **OK**
11. Expand *Physical Capacity*
12. Expand to the lowest level of *Data Center*
13. Drag *Data Center* to the Reporting palate.
14. Expand *Capacity (MB)*
15. Drag *Capacity (MB)* to the Reporting palate.
16. Drag *Used Capacity (MB)* to the Reporting palate.
17. Run the report by clicking  and selecting an output type.

Result

A report similar to the following is created:

	Data Center	Capacity (MB)	Used Capacity (MB)
	Asia	122,070,096.00	45,708,105.00
	BLR	100,709,506.00	54,982,204.00
	Boulder	22,883,450.00	12,011,075.00
	DC01	1,707,024,715.00	1,407,609,686.00
	DC02	732,370,688.00	732,370,688.00
	DC03	314,598,162.00	65,448,975.00
	DC04	573,573,884.00	282,645,615.00
	DC05	89,245,458.00	62,145,011.00
	DC06	19,455,433,799.00	11,283,487,744.00
	DC08	100,709,506.00	44,950,171.00
	DC10	112,916,718.00	43,346,818.00
	DC14	23,565,735,054.00	17,357,431,924.00
	DC56	137,549,084.00	10,657,793.00
	Europe	743,942,208.00	240,369,325.00
	HIO	9,823,036,853.00	4,216,750,338.00
	London	0.00	0.00
	N/A	9,049,939,023.00	5,887,911,992.00
	RTP	12,386,326,262.00	5,638,948,477.00
	SAC	9,269,642,330.00	6,197,549,437.00
	↑ Top ↑ Page up ↓ Page down ↓ Bottom		

Managing Reports

You can customize a report's output format and delivery, set report properties or schedules, and email reports.



The Reporting feature is available in Cloud Insights Premium Edition.

Customizing a report's output format and delivery

You can customize the format and delivery method of reports.

1. In the Cloud Insights Reporting Portal, mouse over the report you want to customize and click [...].

The screenshot shows a list of four reports on the left:

- Storage Infrastructure Executive Summary (4/29/2019 8:30 PM)
- Virtual Machine Remediation (4/29/2019 8:30 PM) - This item has a blue selection bar and three dots at the end of its row.
- Weekly Compute Consumption (4/29/2019 8:31 PM)
- Weekly Storage Consumption (4/29/2019 8:31 PM)

A context menu is open on the right side of the second report, listing the following options:

- Run as
- Edit report
- Create report view
- View versions
- Properties
- Create a new job
- Copy or move
- Create shortcut
- Share
- Delete

1. Click **Properties > Schedule**

< Back

Create schedule

Period

Start	2018-04-06	1:49 PM
End	2018-07-06	1:49 PM
<input type="checkbox"/> No end date		

Run every 1 week(s)

On day(s) M T W T F S S

Daily time interval

Options

Format HTML >

Delivery Save >

Prompts Set values >

Languages English (United States) >

1. You can set the following options:

- Schedule when you want reports to run.
- Format the report output.
- Delivery: print, save, or email the report.
- Languages define the language the report is delivered in.

2. Click **Create** to produce the report using the selections you made.

Copying a report to the clipboard

Use this process to copy a report to the clipboard.

1. In the toolbar, click **[+]**
2. Click **Report**
3. Click the **Pages** icon . The Report icon **Report** is displayed

4. Left click the Report icon. Report options are displayed.

5. Click **Copy Report to Clipboard**.

Opening reports (xml) from the clipboard

You can open a report specification that was previously copied to the clipboard.

About this task

You enter the Reporting user interface by creating a new report or opening an existing report

1. In the toolbar, click **[+]**

2. Click **Report**

3. Click the **Pages** icon  . The Report icon  **Report** is displayed

4. Left click the Report icon. Report options are displayed.

5. Click **Open report from clipboard**.

Creating Custom Reports

You can use the report authoring tools to create custom reports. After creating reports, you can save them and run them on a regular schedule. The results of reports can be automatically sent by email to yourself and others.



The Reporting feature is available in Cloud Insights [Premium Edition](#).

The examples in this section show the following process, which can be used for any of the Cloud Insights Reporting data models:

- Identifying a question to be answered with a report
- Determining the data needed to support the results
- Selecting data elements for the report

Before designing your custom report, you need to complete some prerequisite tasks. If you do not complete these, reports could be inaccurate or incomplete.

For example, if you do not finish the device identification process, your capacity reports will not be accurate. Or, if you do not finish setting annotations (such as tiers, business units, and data centers), your custom reports might not accurately report data across your domain or might show "N/A" for some data points.

Before you design your reports, complete the following tasks:

- Configure all [data collectors](#) properly.
- Enter annotations (such as tiers, data centers, and business units) on devices and resources in your environment. It is beneficial to have annotations stable before generating reports, because Cloud Insights Reporting collects historical information.

Report Creation Process

The process of creating custom (also called "ad hoc") reports involves several tasks:

- Plan the results of your report.
- Identify data to support your results.
- Select the data model (for example, Chargeback data model, Inventory data model, and so on) that contains the data.
- Select data elements for the report.
- Optionally format, sort, and filter report results.

Planning the Results of Your Custom Report

Before you open the report authoring tools, you might want to plan the results you want from the report. With report authoring tools, you can create reports easily and might not need a great deal of planning; however, it is a good idea to get a sense from the report requestor about the report requirements.

- Identify the exact question you want to answer. For example:
 - How much capacity do I have left?
 - What are the chargeback costs per business unit?
 - What is the capacity by tier to ensure that business units are aligned at the proper tier of storage?
 - How can I forecast power and cooling requirements? (Add customized metadata by adding annotations to resources.)
- Identify the data elements that you need to support the answer.
- Identify the relationships between data that you want to see in the answer. Do not include illogical relationships in your question, for example, “I want to see the ports that relate to capacity.”
- Identify any calculations needed on data.
- Determine what types of filtering are needed to limit the results.
- Determine if you need to use current or historical data.
- Determine if you need to set access privileges on reports to limit the data to specific audiences.
- Identify how the report will be distributed. For example, should it be emailed on a set schedule or included in the Team content folder area?
- Determine who will maintain the report. This might affect the complexity of the design.
- Create a mockup of the report.

Tips for designing reports

Several tips might be helpful when you are designing reports.

- Determine whether you need to use current or historical data.

Most reports only need to report on the latest data available in the Cloud Insights.

- Cloud Insights Reporting provides historical information on capacity and performance, but not on inventory.
- Everybody sees all data; however, you might need to limit data to specific audiences.

To segment the information for different users, you can create reports and set access permissions on them.

Reporting data models

Cloud Insights includes several data models from which you can either select predefined reports or create your own custom report.

Each data model contains a simple data mart and an advanced data mart:

- The simple data mart provides quick access to the most commonly used data elements and includes only the last snapshot of Data Warehouse data; it does not include historical data.
- The advanced data mart provides all values and details available from the simple data mart and includes access to historical data values.

Capacity data model

Enables you to answer questions about storage capacity, file system utilization, internal volume capacity, port capacity, qtree capacity, and virtual machine (VM) capacity. The Capacity data model is a container for several capacity data models. You can create reports answering various types of questions using this data model:

Storage and Storage Pool Capacity data model

Enables you to answer questions about storage capacity resource planning, including storage and storage pools, and includes both physical and virtual storage pool data. This simple data model can help you answer questions related to capacity on the floor and the capacity usage of storage pools by tier and data center over time.

If you are new to capacity reporting, you should start with this data model because it is a simpler, targeted data model. You can answer questions similar to the following using this data model:

- What is the projected date for reaching the capacity threshold of 80% of my physical storage?
- What is the physical storage capacity on an array for a given tier?
- What is my storage capacity by manufacturer and family as well as by data center?
- What is the storage utilization trend on an array for all of the tiers?
- What are my top 10 storage systems with the highest utilization?
- What is the storage utilization trend of the storage pools?
- How much capacity is already allocated?
- What capacity is available for allocation?

File System Utilization data model

This data model provides visibility about capacity utilization by hosts at the file system level. Administrators can determine allocated and used capacity per file system, determine the type of file system, and identify trending statistics by file system type. You can answer the following questions using this data model:

- What is the size of the file system?
- Where is the data kept and how is it accessed, for example, local or SAN?
- What are the historical trends for the file system capacity? Then, based on this, what can we anticipate for future needs?

Internal Volume Capacity data model

Enables you to answer questions about internal volume used capacity, allocated capacity, and capacity usage

over time:

- Which internal volumes have a utilization higher than a predefined threshold?
- Which internal volumes are in danger of running out of capacity based on a trend?
- What is the used capacity versus the allocated capacity on our internal volumes?

Port Capacity data model

Enables you to answer questions about switch port connectivity, port status, and port speed over time. You can answer questions similar the following to help you plan for purchases of new switches:

How can I create a port consumption forecast that predicts resource (port) availability (according to data center, switch vendor and port speed)?

- Which ports are likely to run out of capacity, providing data speed, data center, vendor and number of Host and storage ports?
- What are the switch port capacity trends over time?
- What are the port speeds?
- What type of port capacity is needed and which organization is about to run out of a certain port type or vendor?
- What is the optimal time to purchase that capacity and make it available?

Qtree Capacity data model

Enables you to trend qtree utilization (with data such as used versus allocated capacity) over time. You can view the information by different dimensions—for example, by business entity, application, tier, and service level. You can answer the following questions using this data model:

- What is the used capacity for qtrees versus the limits set per application or business entity?
- What are the trends of our used and free capacity so that we can do capacity planning?
- Which business entities are using the most capacity?
- Which applications consume the most capacity?

VM Capacity data model

Enables you to report your virtual environment and its capacity usage. This data model lets you report on changes in capacity usage over time for VMs and data stores. The data model also provides thin provisioning and virtual machine chargeback data.

- How can I determine capacity chargeback based on capacity provisioned to VMs and data stores?
- What capacity is not used by VMs and which portion of unused is free, orphaned, or other?
- What do we need to purchase based on consumption trends?
- What are my storage efficiency savings achieved by using storage thin provisioning and deduplication technologies?

Capacities in the VM Capacity data model are taken from virtual disks (VMDKs). This means that the provisioned size of a VM using the VM Capacity data model is the size of its virtual disks. This is different from the provisioned capacity in the Virtual Machines view in Cloud Insights, which shows the provisioned size for the VM itself.

Volume Capacity data model

Enables you to analyze all aspects of the volumes in your environment and organize data by vendor, model, tier, service level, and data center.

You can view the capacity related to orphaned volumes, unused volumes, and protection volumes (used for replication). You can also see different volume technologies (iSCSI or FC), and compare virtual volumes to non-virtual volumes for array virtualization issues.

You can answer questions similar to the following with this data model:

- Which volumes have a utilization higher than a predefined threshold?
- What is the trend in my data center for orphan volume capacity?
- How much of my data center capacity is virtualized or thin provisioned?
- How much of my data center capacity must be reserved for replication?

Chargeback data model

Enables you to answer questions about used capacity and allocated capacity on storage resources (volumes, internal volumes, and qtrees). This data model provides storage capacity chargeback and accountability information by hosts, application, and business entities, and includes both current and historical data. Report data can be categorized by service level and storage tier.

You can use this data model to generate chargeback reports by finding the amount of capacity that is used by a business entity. This data model enables you to create unified reporting of multiple protocols (including NAS, SAN, FC, and iSCSI).

- For storage without internal volumes, chargeback reports show chargeback by volumes.
- For storage with internal volumes:
 - If business entities are assigned to volumes, chargeback reports show chargeback by volumes.
 - If business entities are not assigned to volumes but assigned to qtrees, chargeback reports show chargeback by qtrees.
 - If business entities are not assigned to volumes and not assigned to qtrees, chargeback reports show the internal volume.
 - The decision whether to show chargeback by volume, qtree or internal volume is made per each internal volume, so it is possible for different internal volumes in the same storage pool to show chargeback at different levels.

Capacity facts are purged after a default time interval. For details, see Data Warehouse processes.

Reports using the Chargeback data model might display different values than reports using the Storage Capacity data model.

- For storage arrays that are not NetApp storage systems, the data from both data models is the same.
- For NetApp and Celerra storage systems, the Chargeback data model uses a single layer (of volumes, internal volumes, or qtrees) to base its charges, while the Storage Capacity data model uses multiple layers (of volumes and internal volumes) to base its charges.

Inventory data model

Enables you to answer questions about inventory resources including hosts, storage systems, switches, disks,

tapes, qtrees, quotas, virtual machines and servers, and generic devices. The Inventory data model includes several submarts that enable you to view information about replications, FC paths, iSCSI paths, NFS paths, and violations. The Inventory data model does not include historical data. Questions you can answer with this data

- What assets do I have and where are they?
- Who is using the assets?
- What types of devices do I have and what are components of those devices?
- How many hosts per OS do I have and how many ports exist on those hosts?
- What storage arrays per vendor exist in each data center?
- How many switches per vendor do I have in each data center?
- How many ports are not licensed?
- What vendor tapes are we using and how many ports exist on each tape? Are all the generic devices identified before we begin working on reports?
- What are the paths between hosts and storage volumes or tapes?
- What are the paths between generic devices and storage volumes or tapes?
- How many violations of each type do I have per data center?
- For each replicated volume, what are the source and target volumes?
- Do I have any firmware incompatibilities or port speed mismatches between Fibre Channel host HBAs and switches?

Performance data model

Enables you to answer questions about performance for volumes, application volumes, internal volumes, switches, applications, VMs, VMDKs, ESX versus VM, hosts, and application nodes. Using this data model, you can create reports that answer several types of performance management questions:

- What volumes or internal volumes have not been used or accessed during a specific period?
- Can we pinpoint any potential misconfiguration for storage for an application (unused)?
- What was the overall access behavior pattern for an application?
- Are tiered volumes assigned appropriately for a given application?
- Could we use cheaper storage for an application currently running without impact to application performance?
- What are the applications that are producing more accesses to currently configured storage?

When you use the switch performance tables, you can obtain the following information:

- Is my host traffic through connected ports balanced?
- Which switches or ports are exhibiting a high number of errors?
- What are the most used switches based on port performance?
- What are the underutilized switches based on port performance?
- What is the host trending throughput based on port performance?
- What is the performance utilization for last X days for one specified host, storage system, tape, or switch?
- Which devices are producing traffic on a specific switch (for example, which devices are responsible for

use of a highly utilized switch)?

- What is the throughput for a specific business unit in our environment?

When you use the disk performance tables, you can obtain the following information:

- What is the throughput for a specified storage pool based on disk performance data?
- What is the highest used storage pool?
- What is the average disk utilization for a specific storage?
- What is the trend of usage for a storage system or storage pool based on disk performance data?
- What is the disk usage trending for a specific storage pool?

When you use VM and VMDK performance tables, you can obtain the following information:

- Is my virtual environment performing optimally?
- Which VMDKs are reporting the highest workloads?
- How can I use the performance reported from VMDs mapped to different datastores to make decisions about re-tiering.

The Performance data model includes information that helps you determine the appropriateness of tiers, storage misconfigurations for applications, and last access times of volumes and internal volumes. This data model provides data such as response times, IOPs, throughput, number of writes pending, and accessed status.

Storage Efficiency data model

Enables you to track the storage efficiency score and potential over time. This data model stores measurements of not only the provisioned capacity, but also the amount that is used or consumed (the physical measurement). For example, when thin provisioning is enabled, Cloud Insights indicates how much capacity is taken from the device. You can also use this model to determine efficiency when deduplication is enabled. You can answer various questions using the Storage Efficiency data mart:

- What is our storage efficiency savings as a result of implementing thin provisioning and deduplication technologies?
- What are the storage savings across data centers?
- Based on historical capacity trends, when do we need to purchase additional storage?
- What would be the capacity gain if we enabled technologies such as thin provisioning and deduplication?
- Regarding storage capacity, am I at risk now?

Data model fact and dimension tables

Each data model includes both fact and dimension tables.

- Fact tables: Contain data that is measured, for example, quantity, raw and usable capacity. Contain foreign keys to dimension tables.
- Dimension tables: Contain descriptive information about facts, for example, data center and business units. A dimension is a structure, often composed of hierarchies, that categorizes data. Dimensional attributes help describe the dimensional values.

Using different or multiple dimension attributes (seen as columns in the reports), you construct reports that

access data for each dimension described in the data model.

Colors used in data model elements

Colors on data model elements have different indications.

- Yellow assets: Represent measurements.
- Non-yellow assets: Represent attributes. These values do not aggregate.

Using multiple data models in one report

Typically, you use one data model per report. However, you can write a report that combines data from multiple data models.

To write a report that combines data from multiple data models, choose one of the data models to use as the base, then write SQL queries to access the data from the additional data marts. You can use the SQL Join feature to combine the data from the different queries into a single query that you can use to write the report.

For example, say you want the current capacity for each storage array and you want to capture custom annotations on the arrays. You could create the report using the Storage Capacity data model. You could use the elements from the Current Capacity and dimension tables and add a separate SQL query to access the annotations information in the Inventory data model. Finally, you could combine the data by linking the Inventory storage data to the Storage Dimension table using the storage name and the join criteria.

How historical data is retained for Reporting

Cloud Insights retains historical data for use in Reporting based on the data marts and granularity of the data, as shown in the following table.

Data mart	Measured object	Granularity	Retention period
Performance marts	Volumes and internal volumes	Hourly	14 days
Performance marts	Volumes and internal volumes	Daily	13 months
Performance marts	Application	Hourly	13 months
Performance marts	Host	Hourly	13 months
Performance marts	Switch performance for port	Hourly	35 days
Performance marts	Switch performance for host, storage, and tape	Hourly	13 months
Performance marts	Storage node	Hourly	14 days
Performance marts	Storage node	Daily	13 months
Performance marts	VM performance	Hourly	35 days
Performance marts	VM performance	Daily	13 months
Performance marts	Hypervisor performance	Hourly	35 days
Performance marts	Hypervisor performance	Daily	13 months

Performance marts	VMDK performance	Hourly	35 days
Performance marts	VMDK performance	Daily	13 months
Performance marts	Disk performance	Hourly	14 days
Performance marts	Disk performance	Daily	13 months
Capacity marts	All (except individual volumes)	Daily	13 months
Capacity marts	All (except individual volumes)	Monthly representative	14 months and beyond
Inventory marts	Individual volumes	Current state	1 day (or until next ETL)

Cloud Insights Data Warehouse Schema Diagrams

This document provides the schema diagrams for the Data Warehouse Database. You can also download a file containing the [schema tables](#).

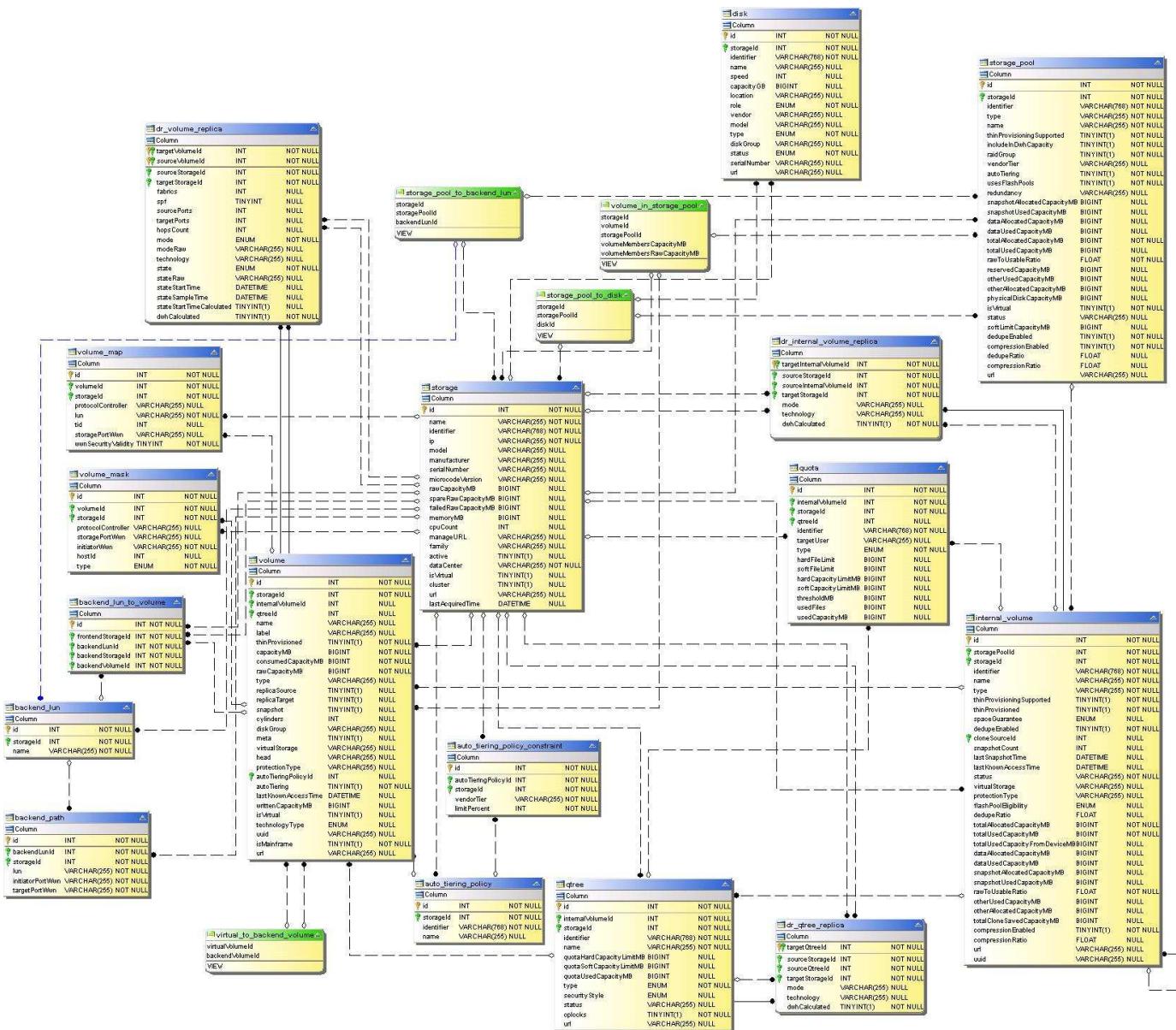


The Reporting feature is available in Cloud Insights [Premium Edition](#).

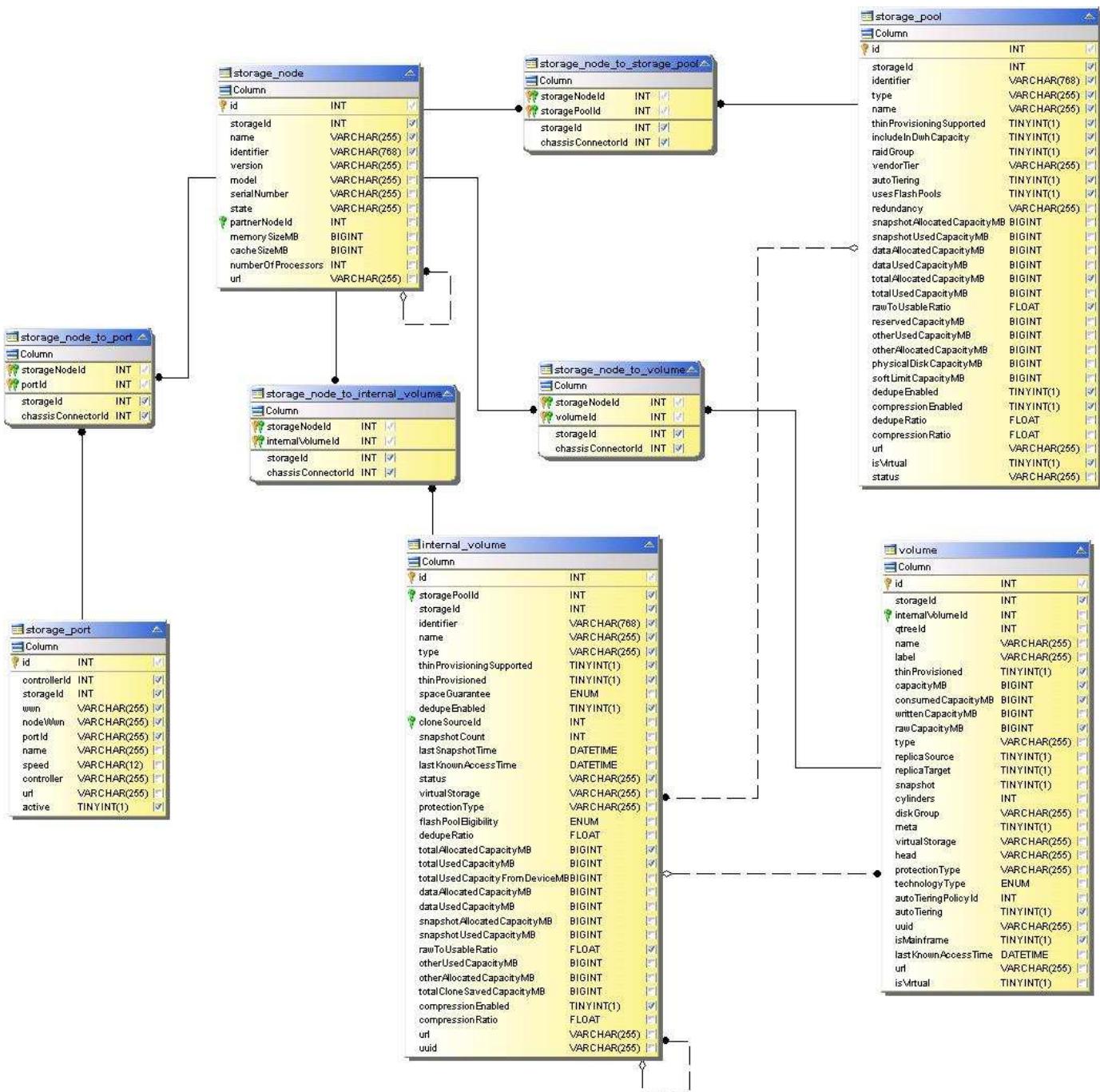
Inventory Datamart

The following images describe the inventory datamart.

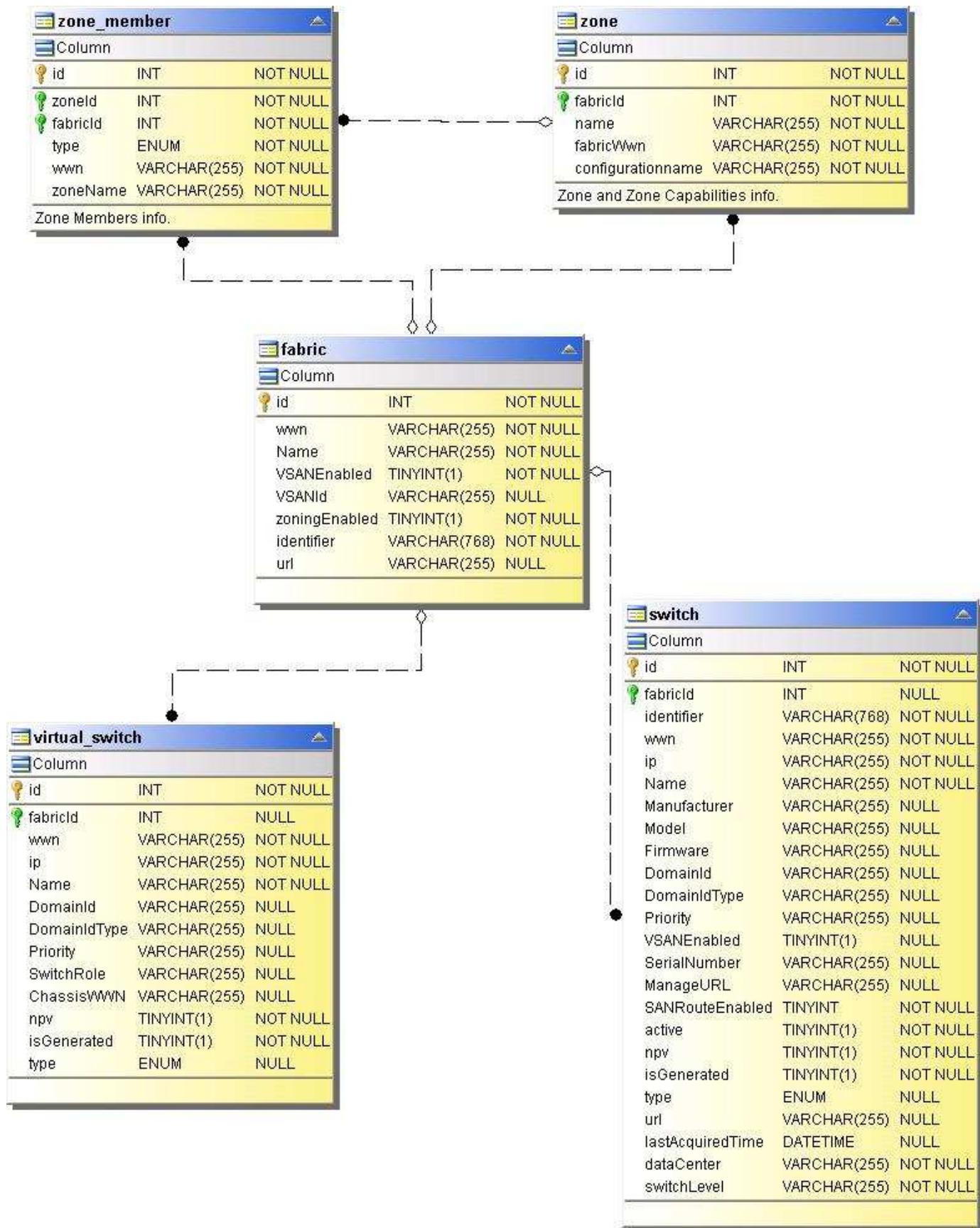
Storage



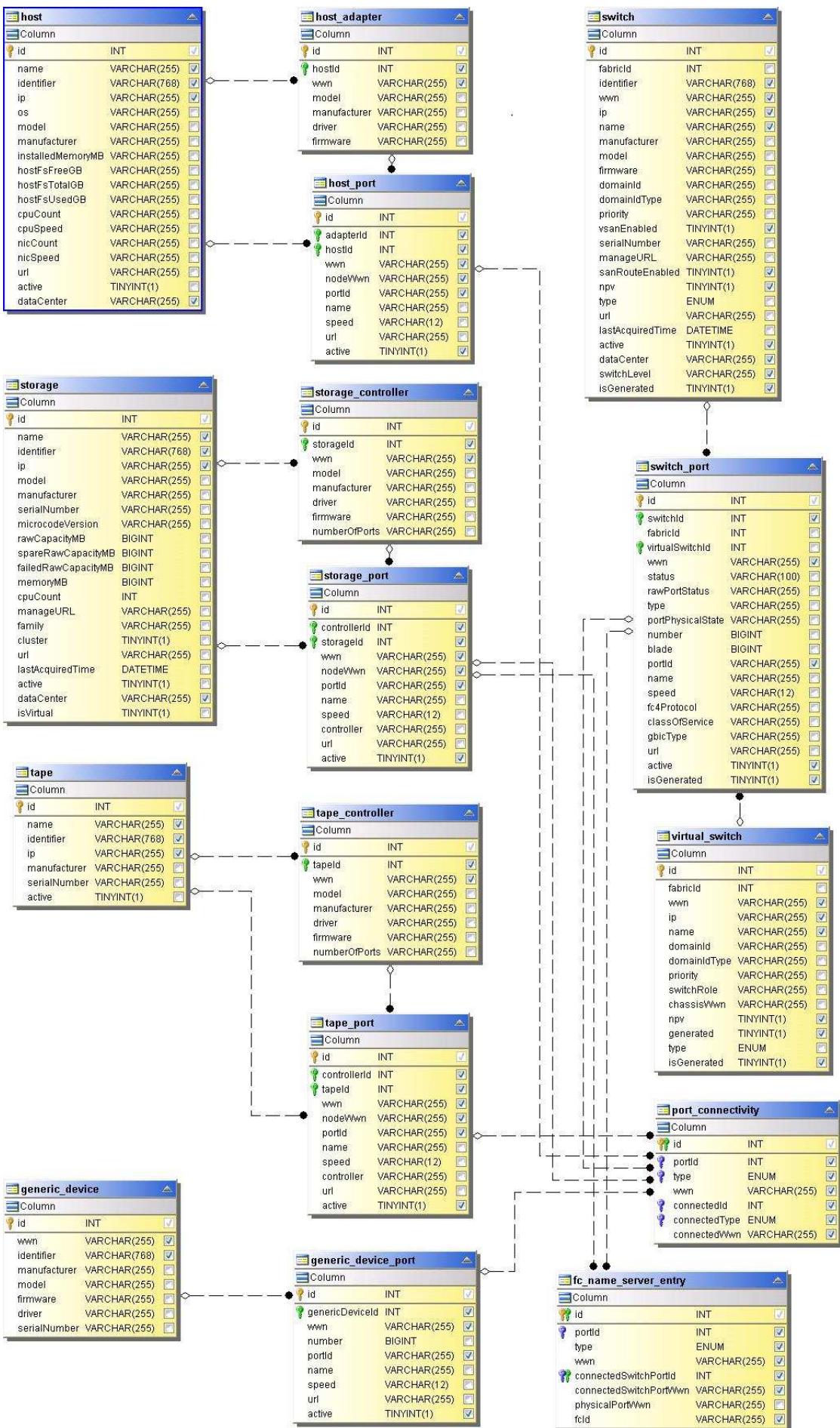
Storage Node

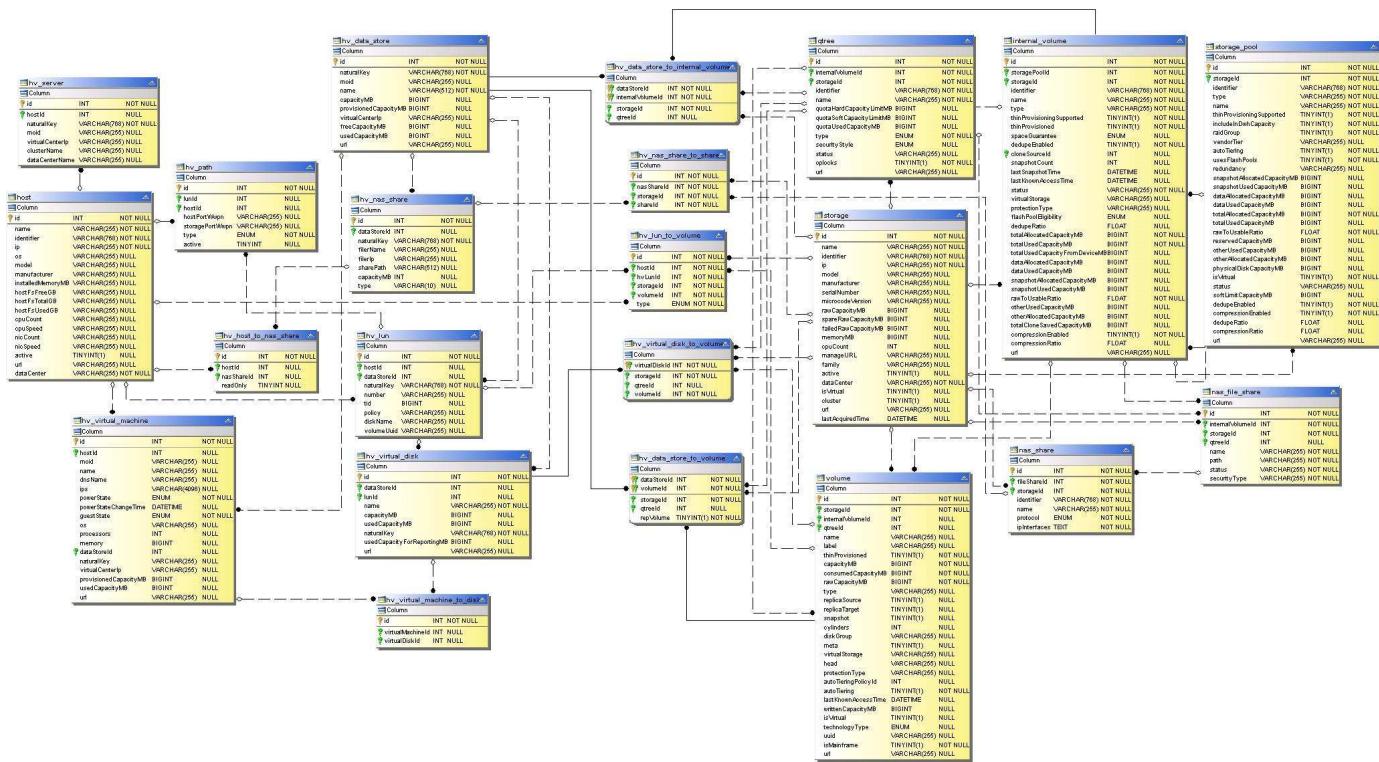


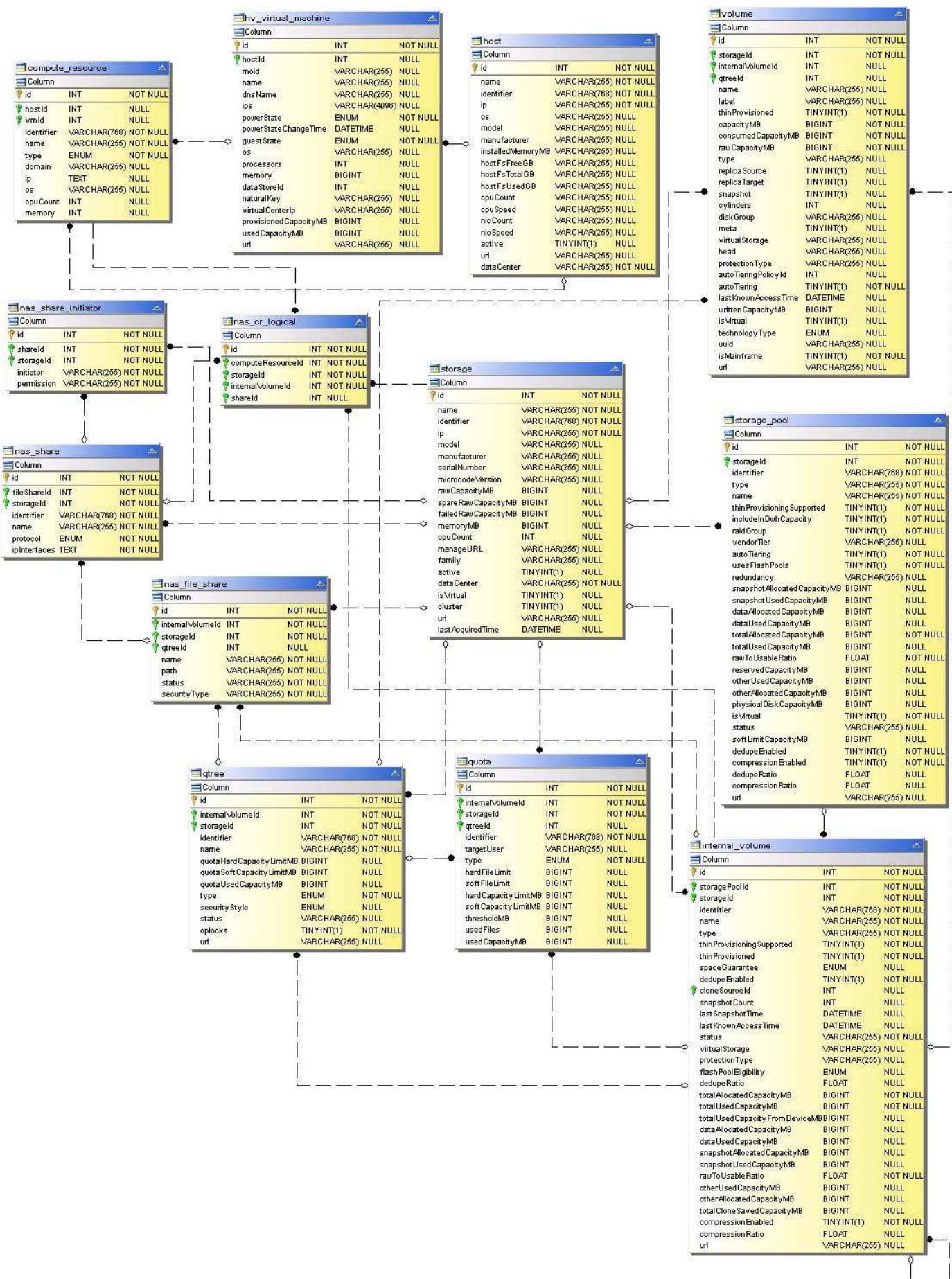
SAN Fabric



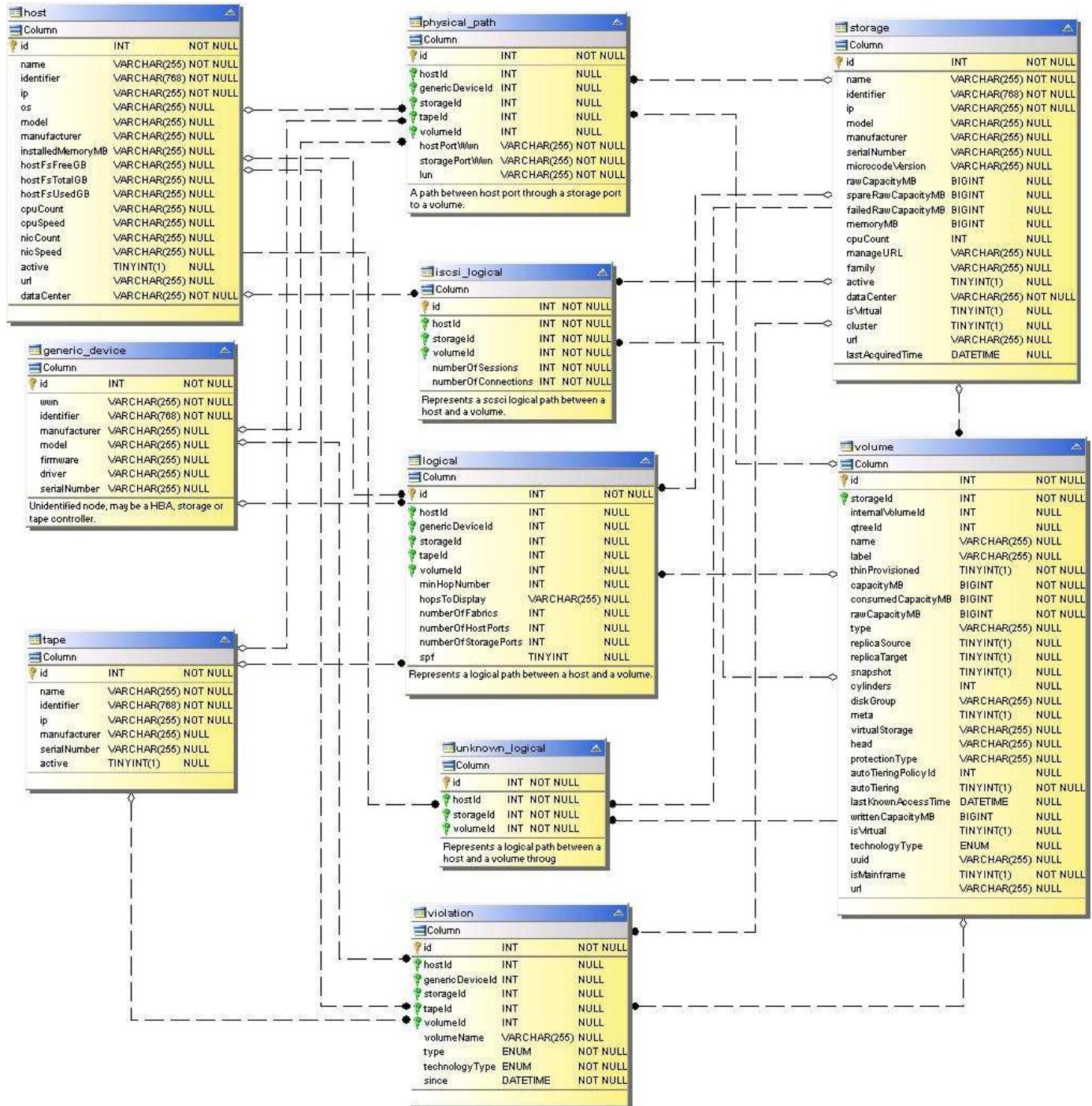
Port Connectivity



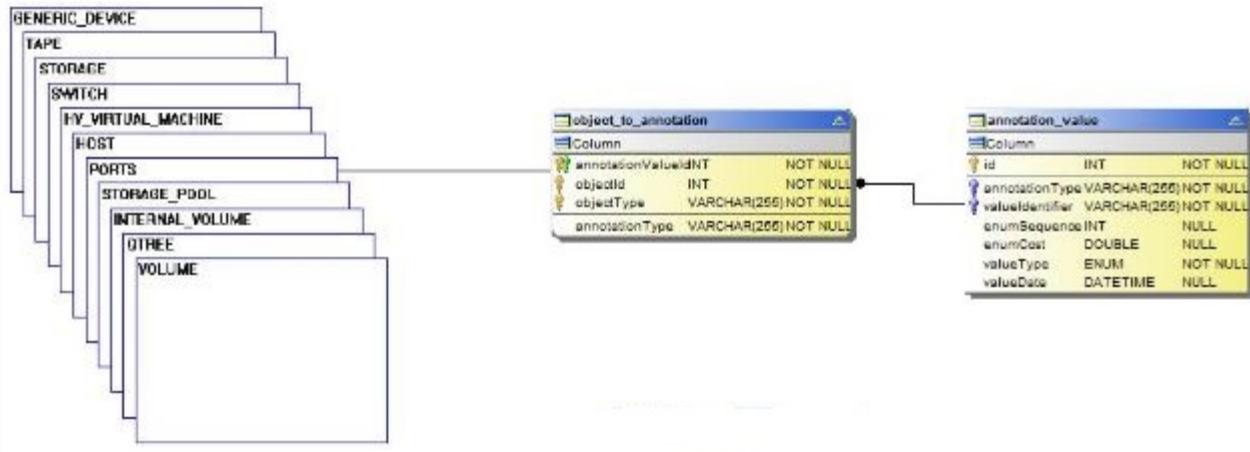




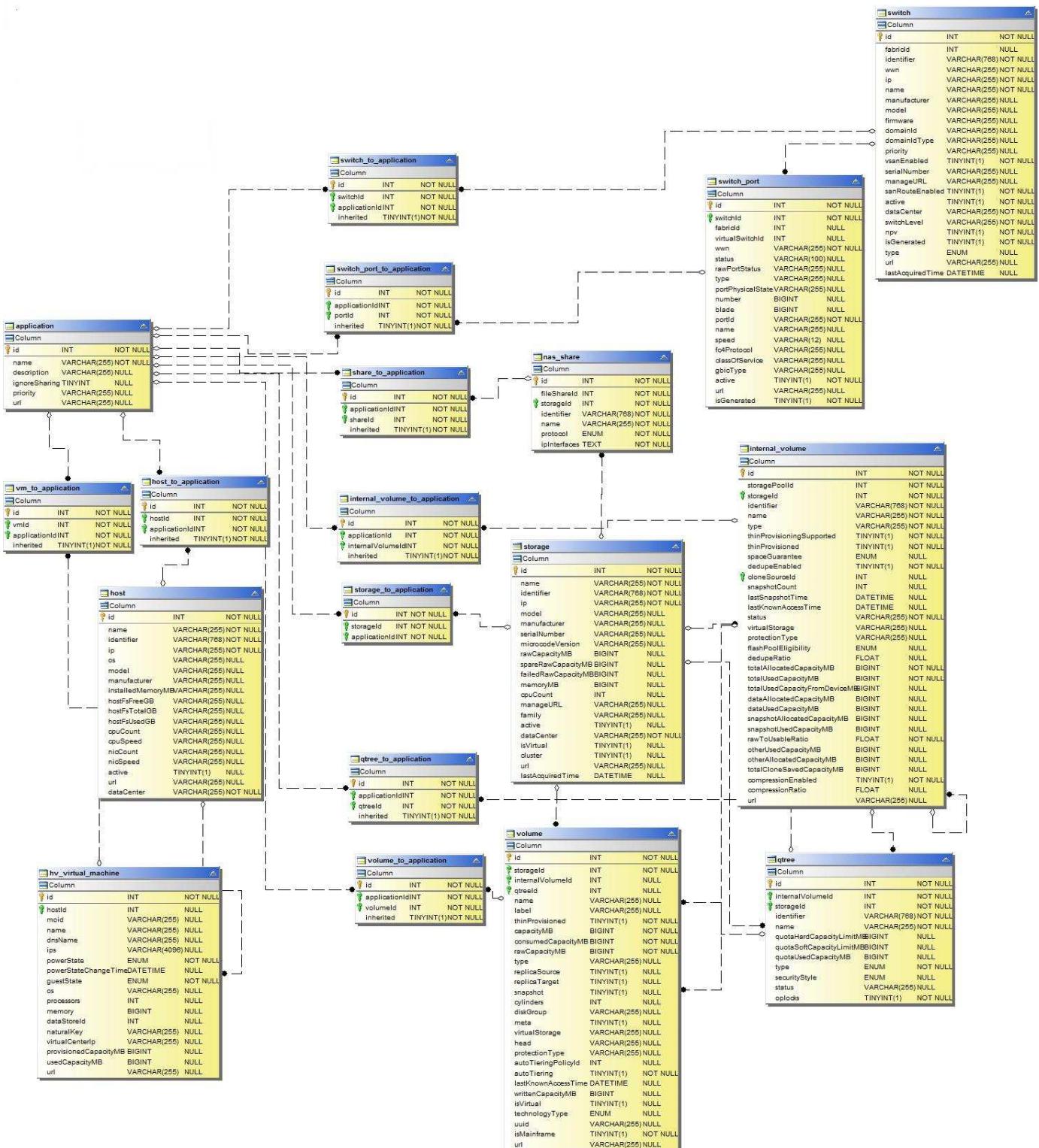
Paths and Violations



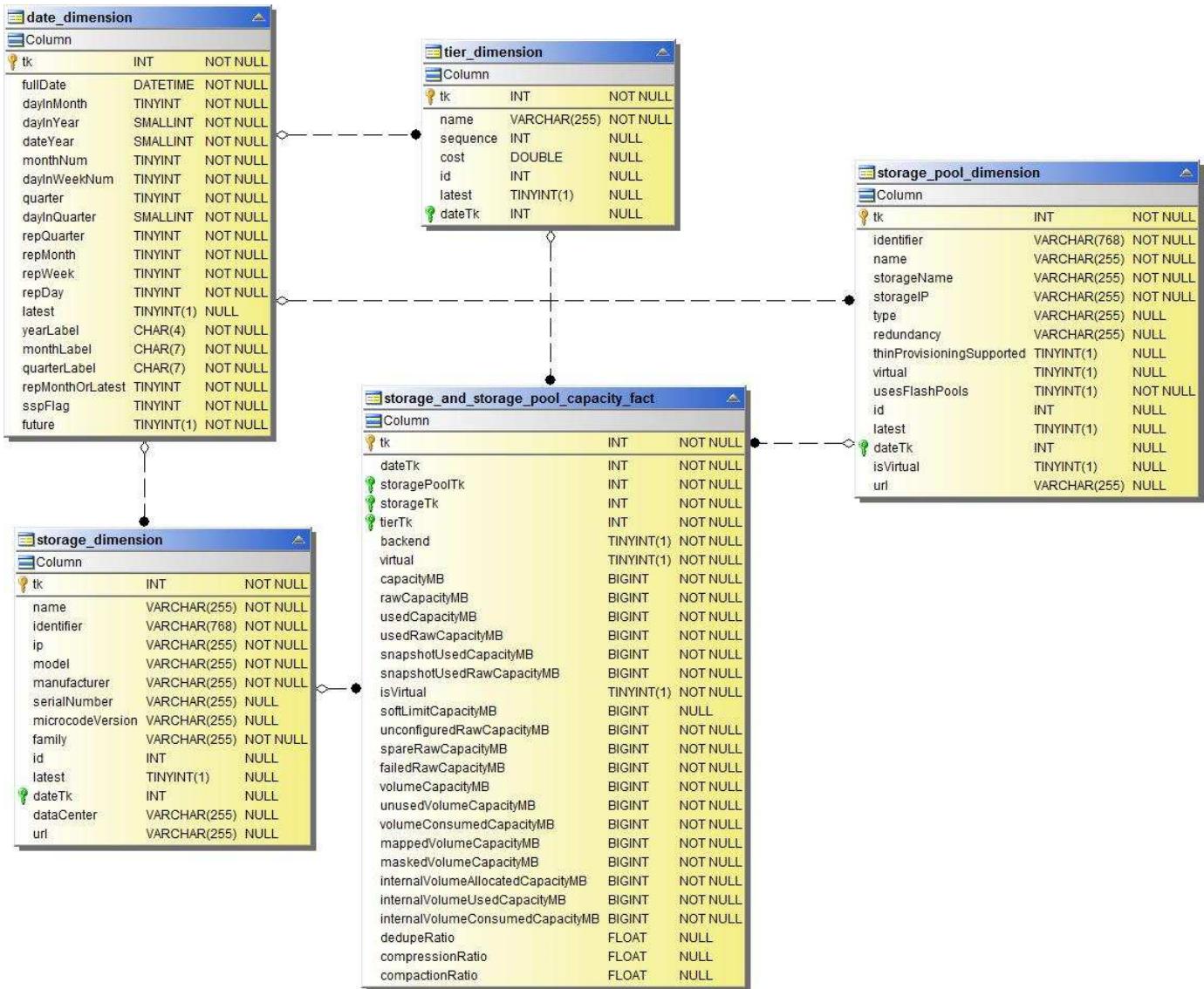
Annotations



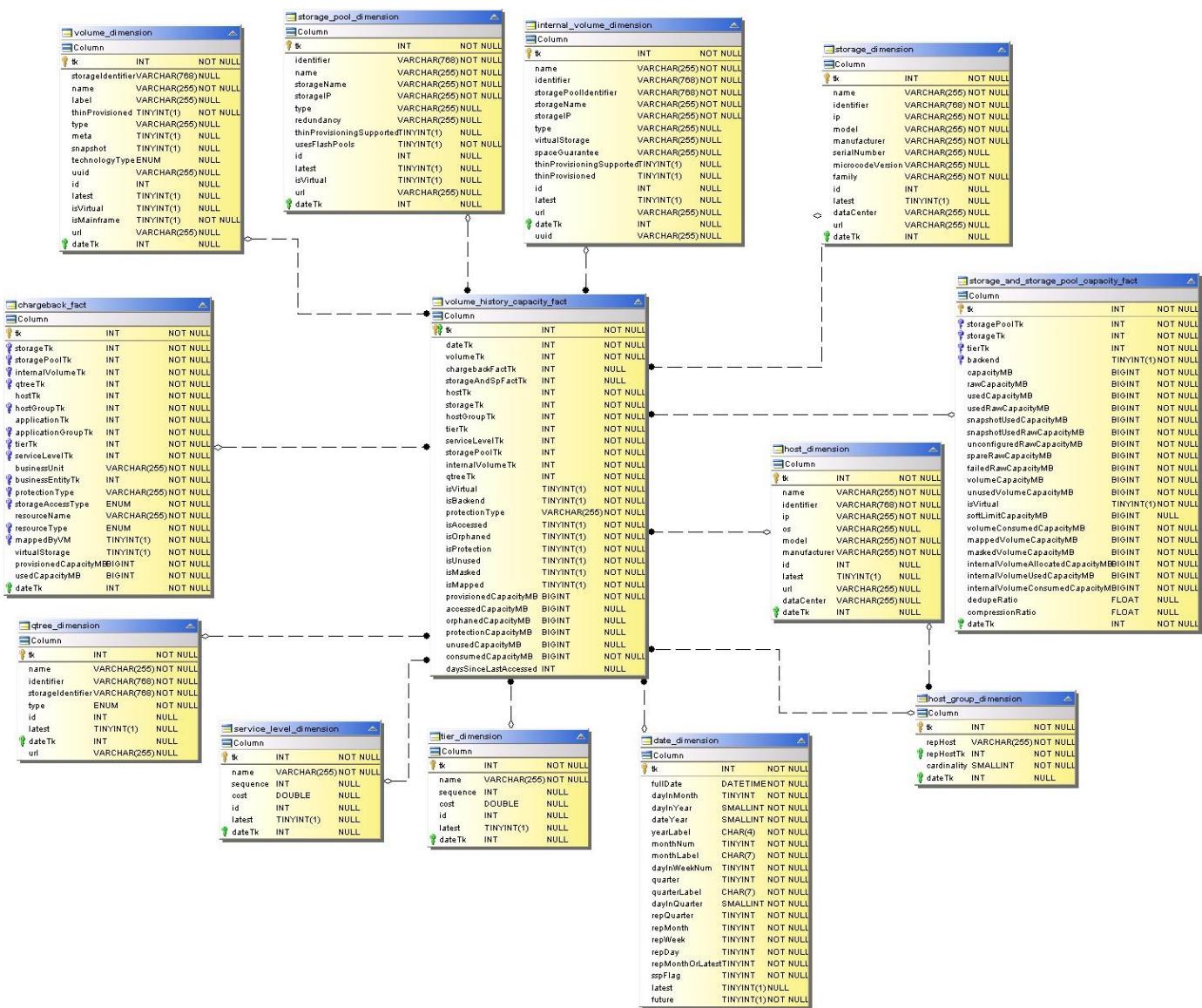
Applications



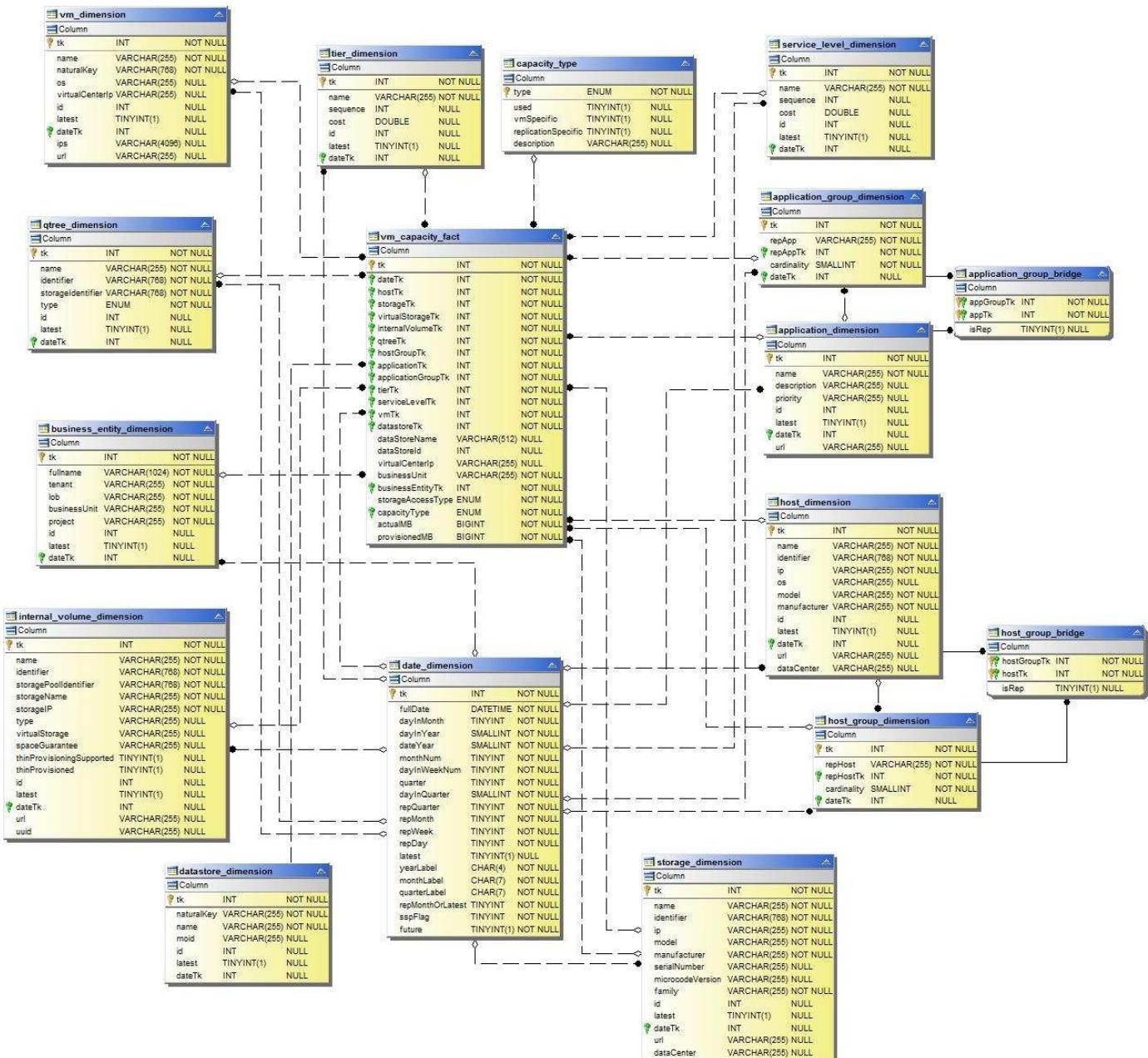
Storage and Storage Pool Capacity



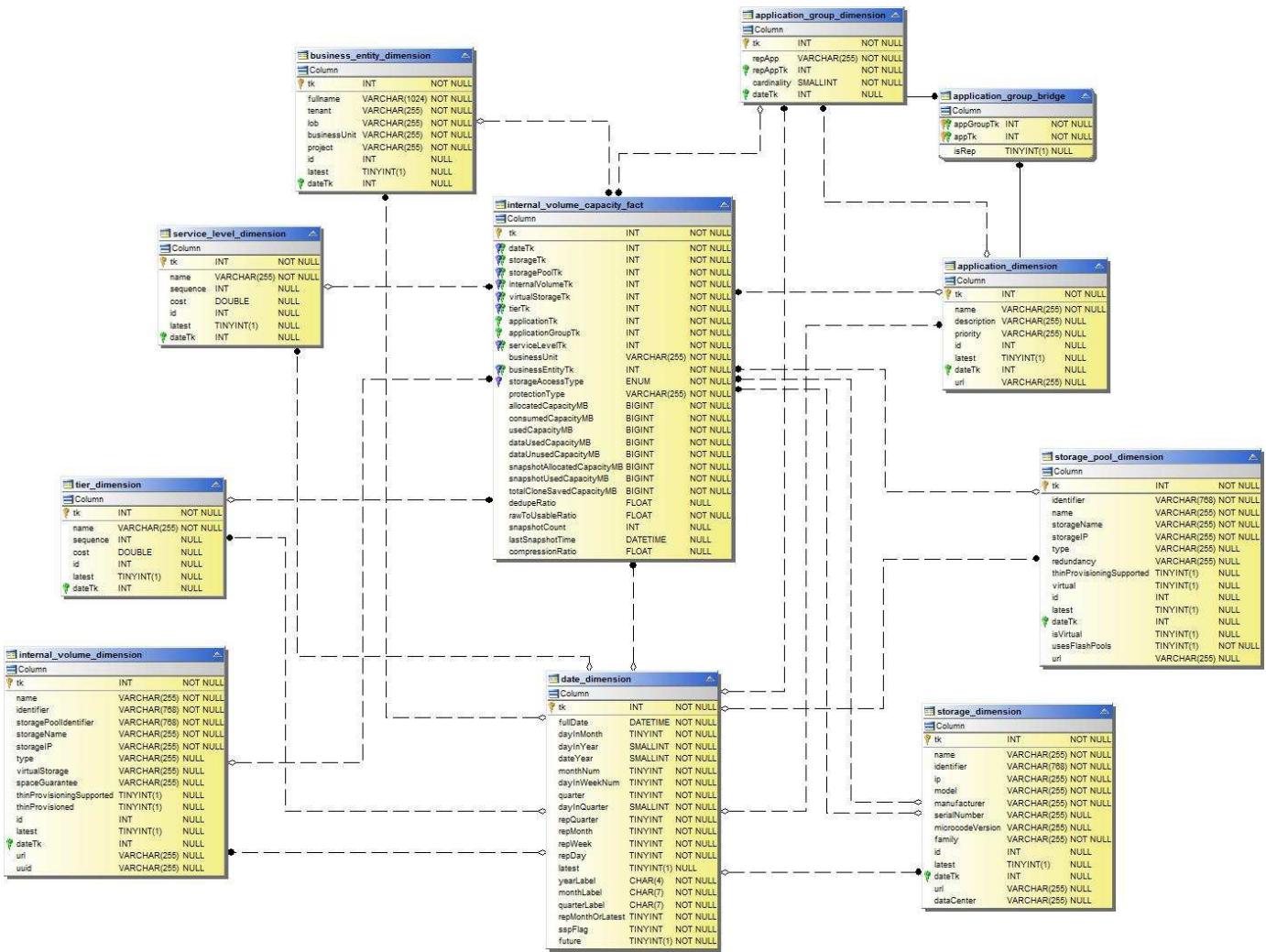
Volume Capacity



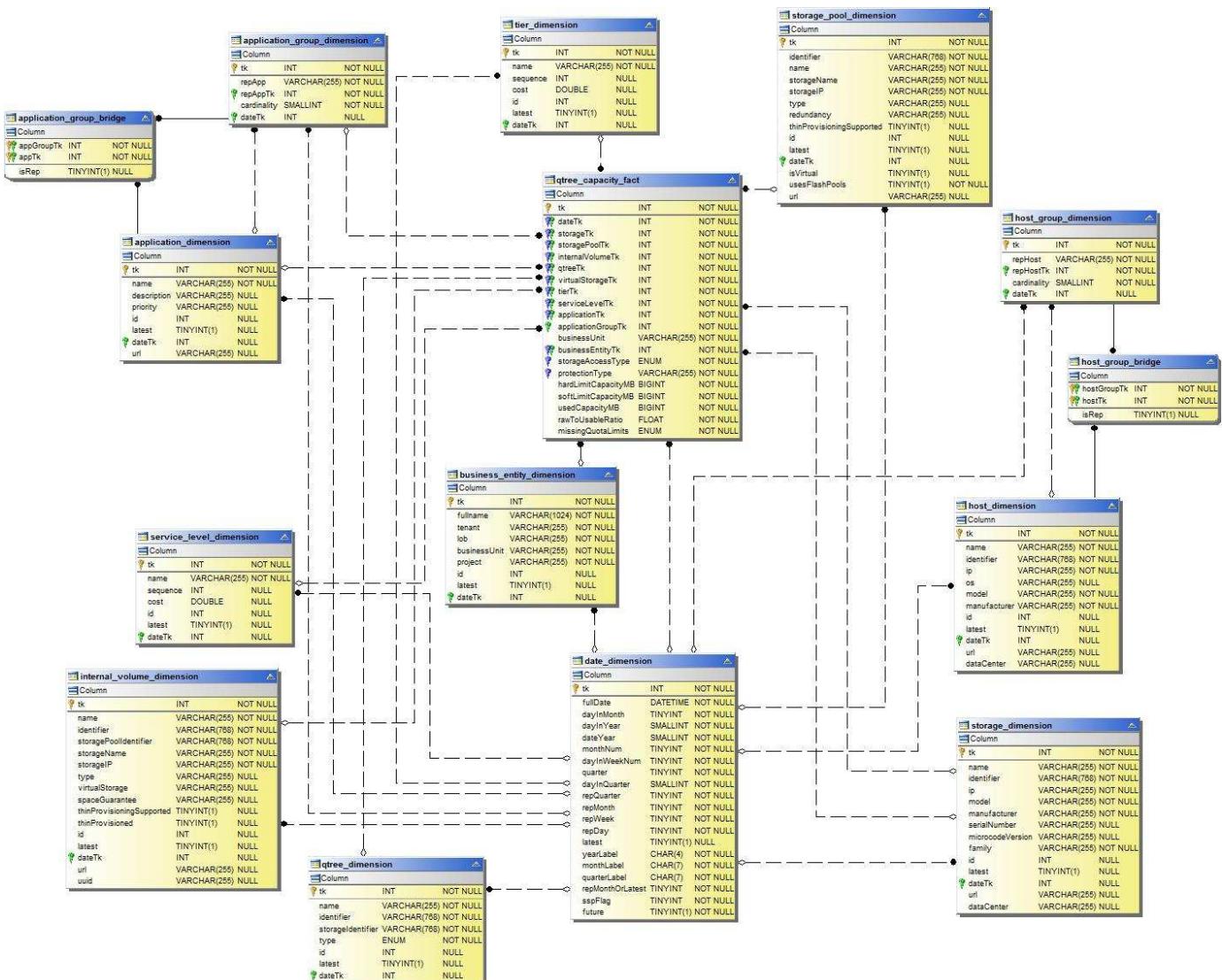
VM Capacity



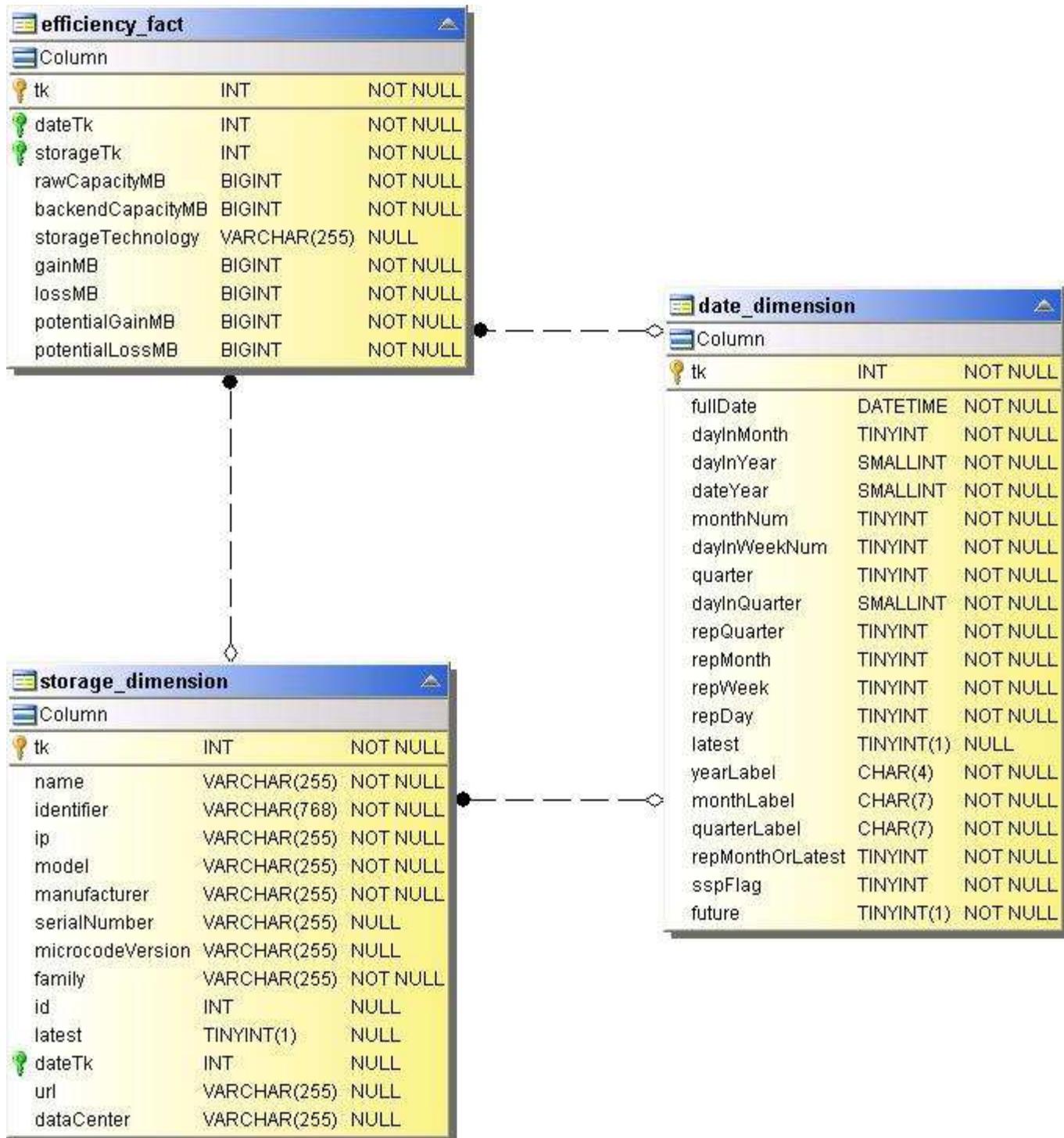
Internal Volume Capacity



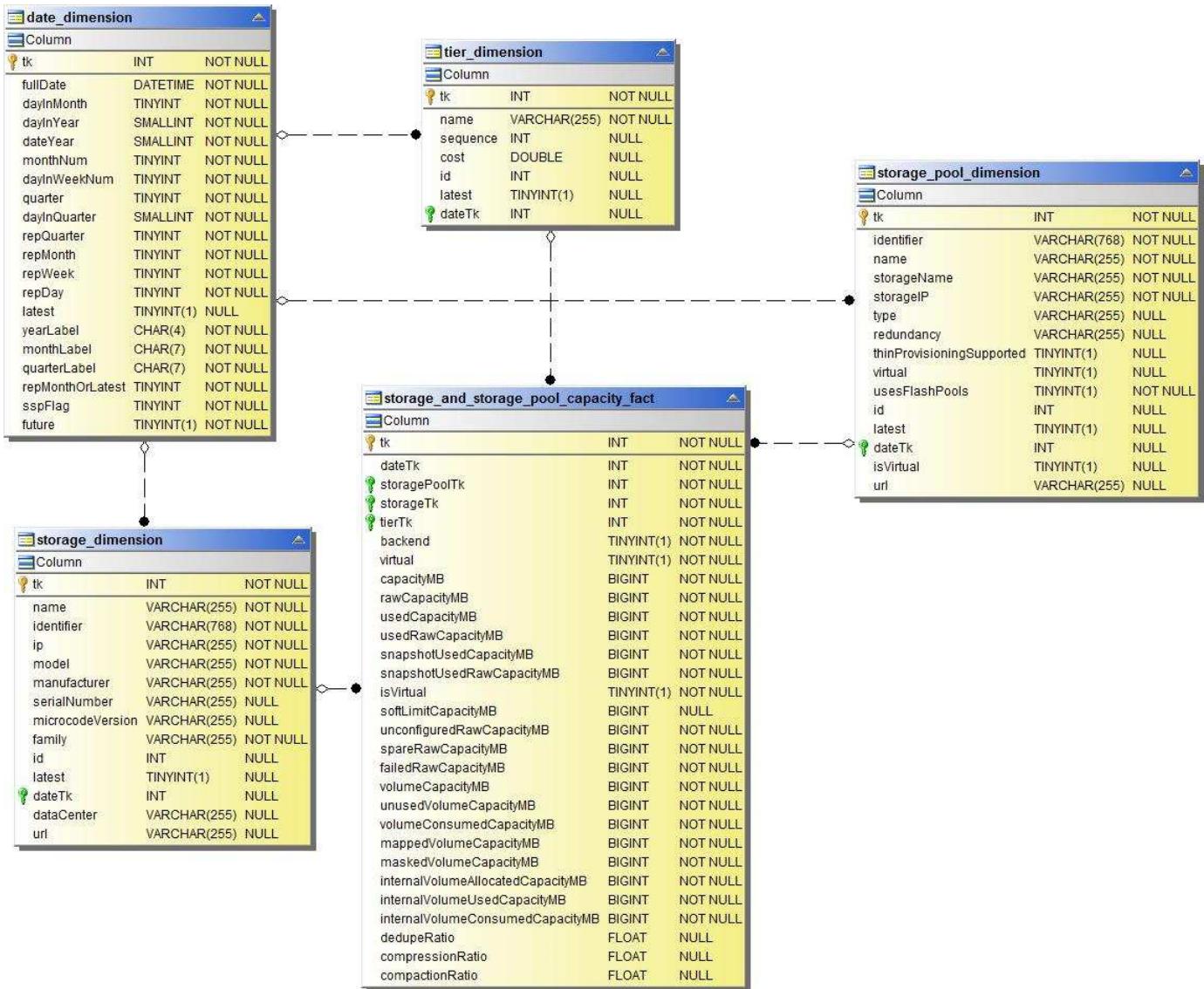
Qtree Capacity



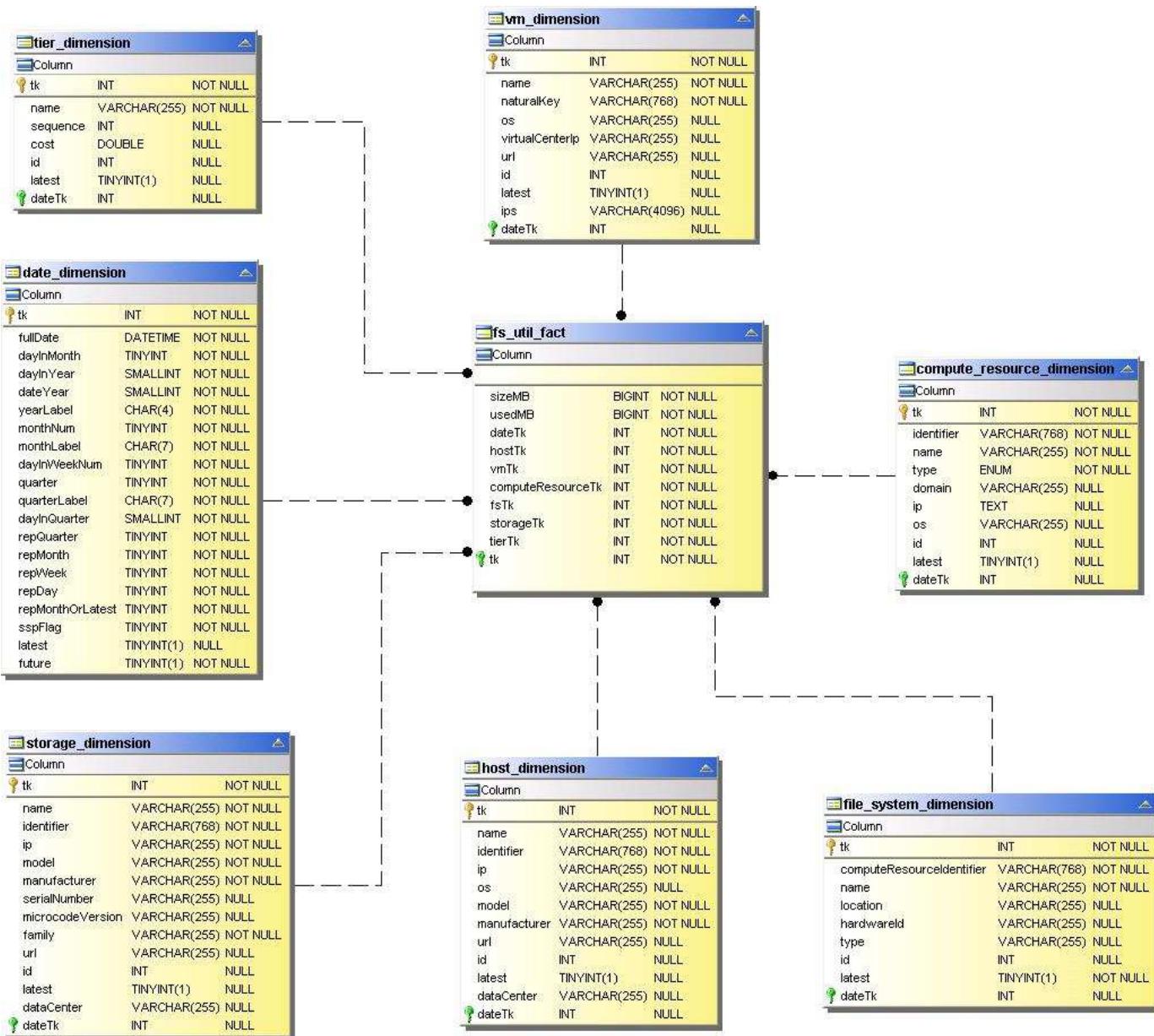
Storage Capacity Efficiency



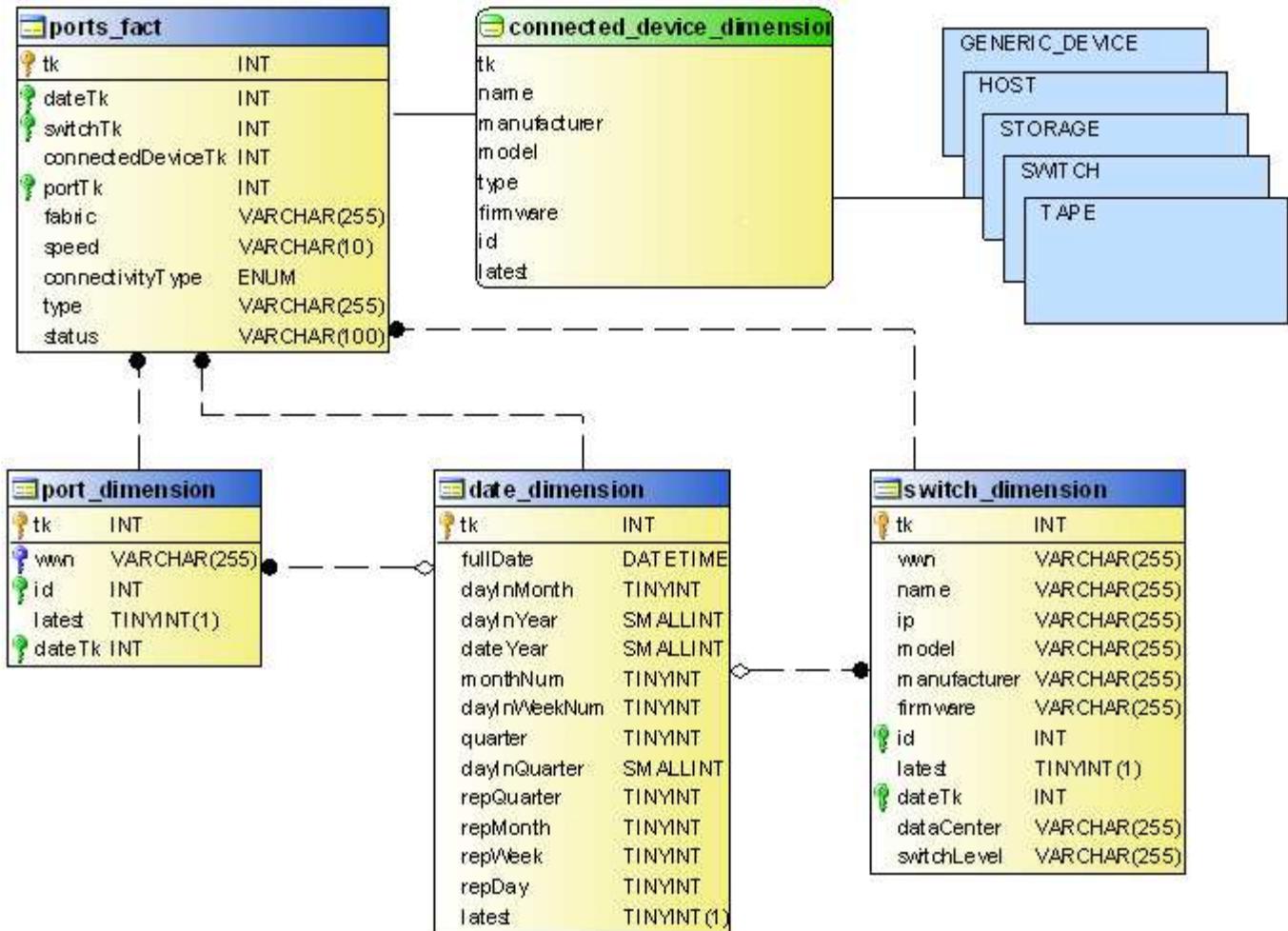
Storage and Storage Pool Capacity



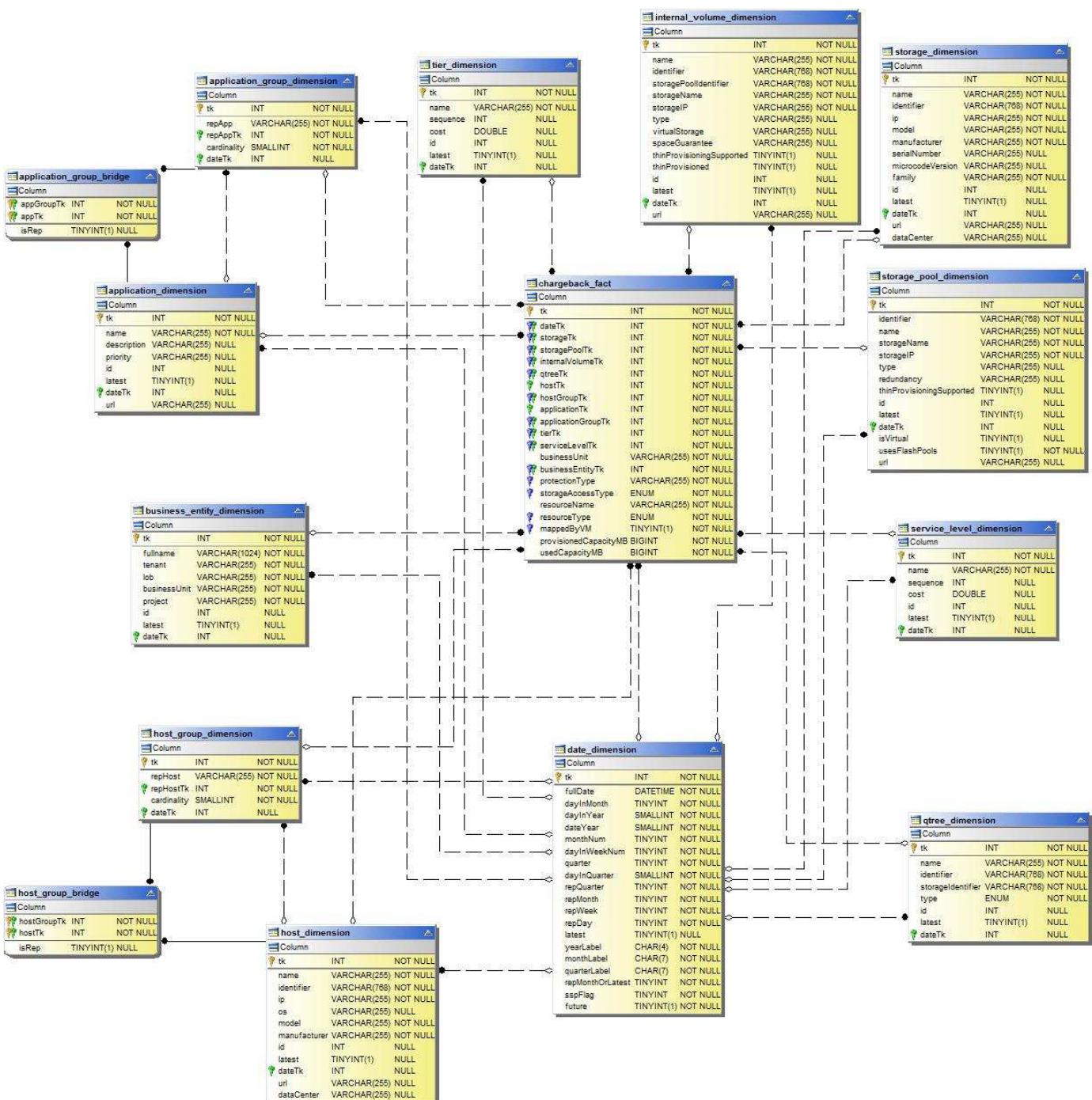
File System Utilization



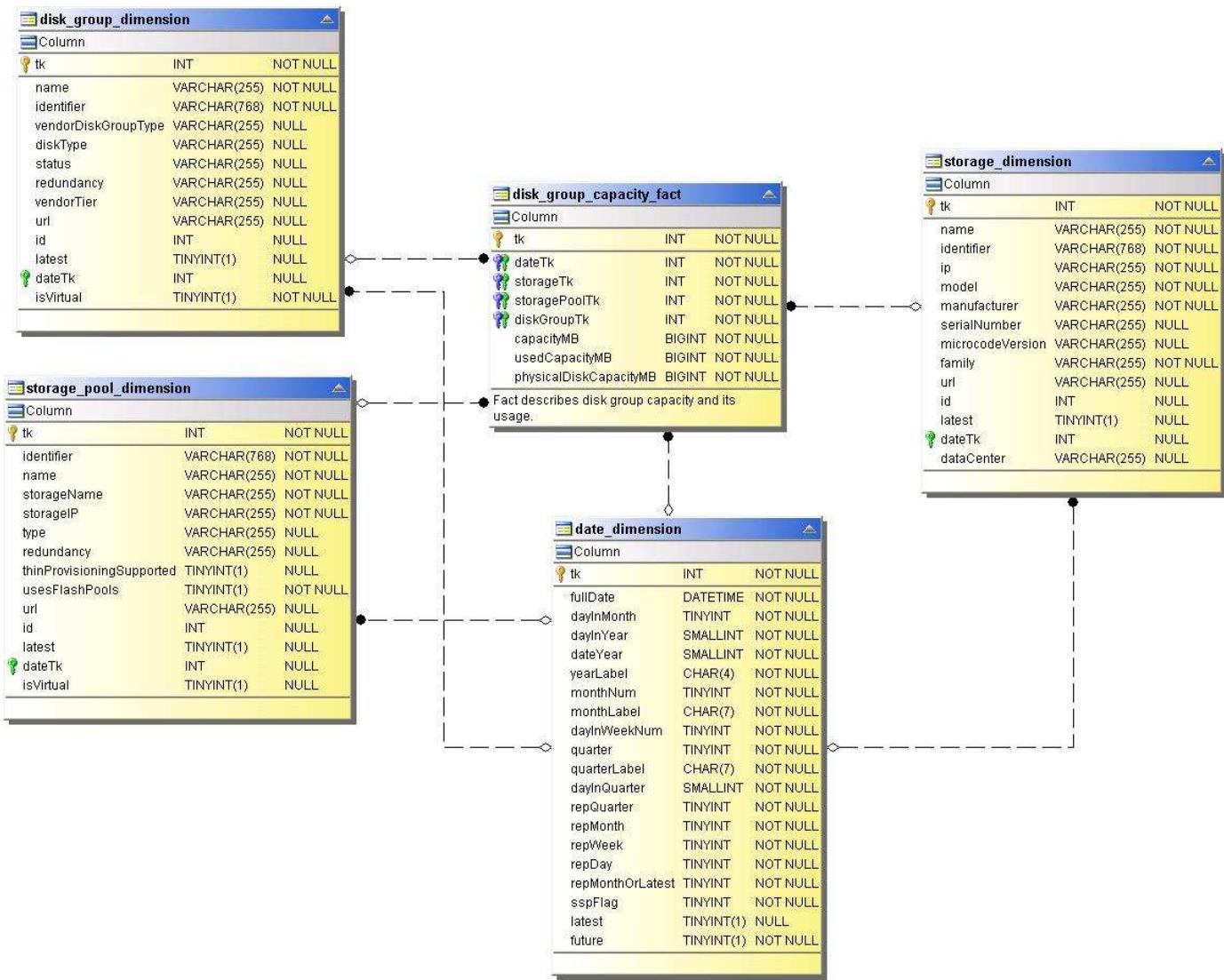
Port Capacity



Chargeback



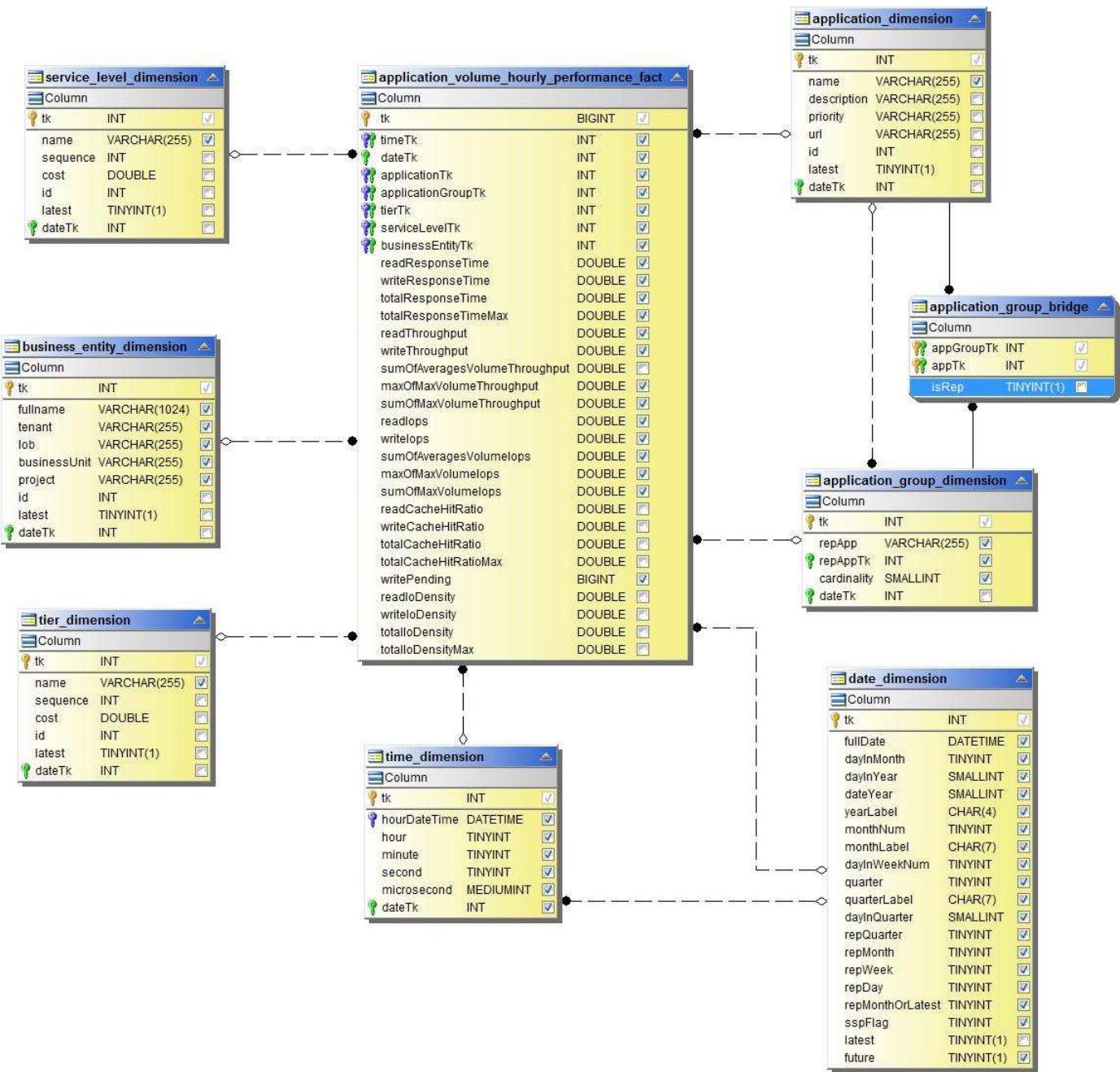
Disk Group Capacity



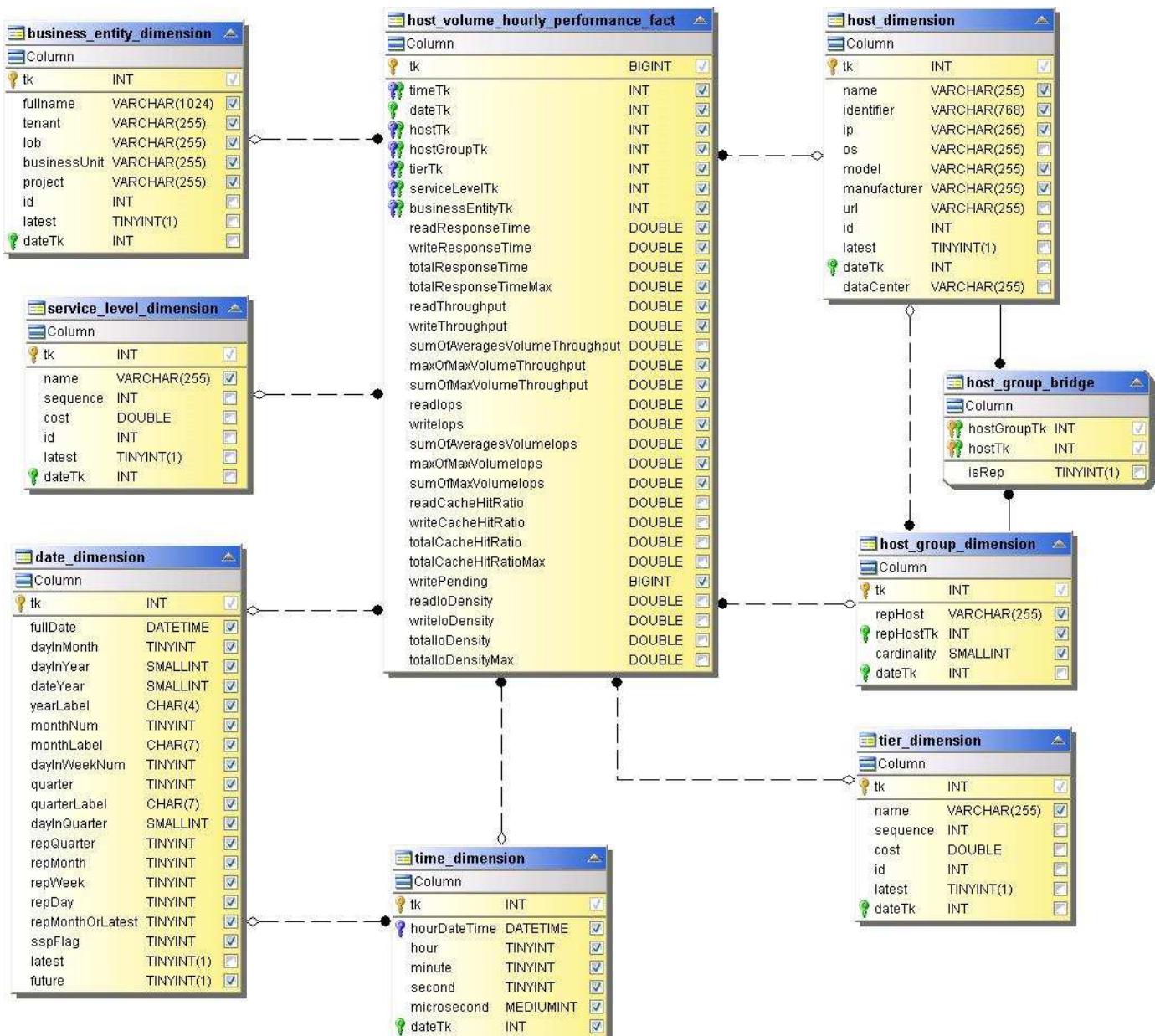
Performance Datamart

The following images describe the performance datamart.

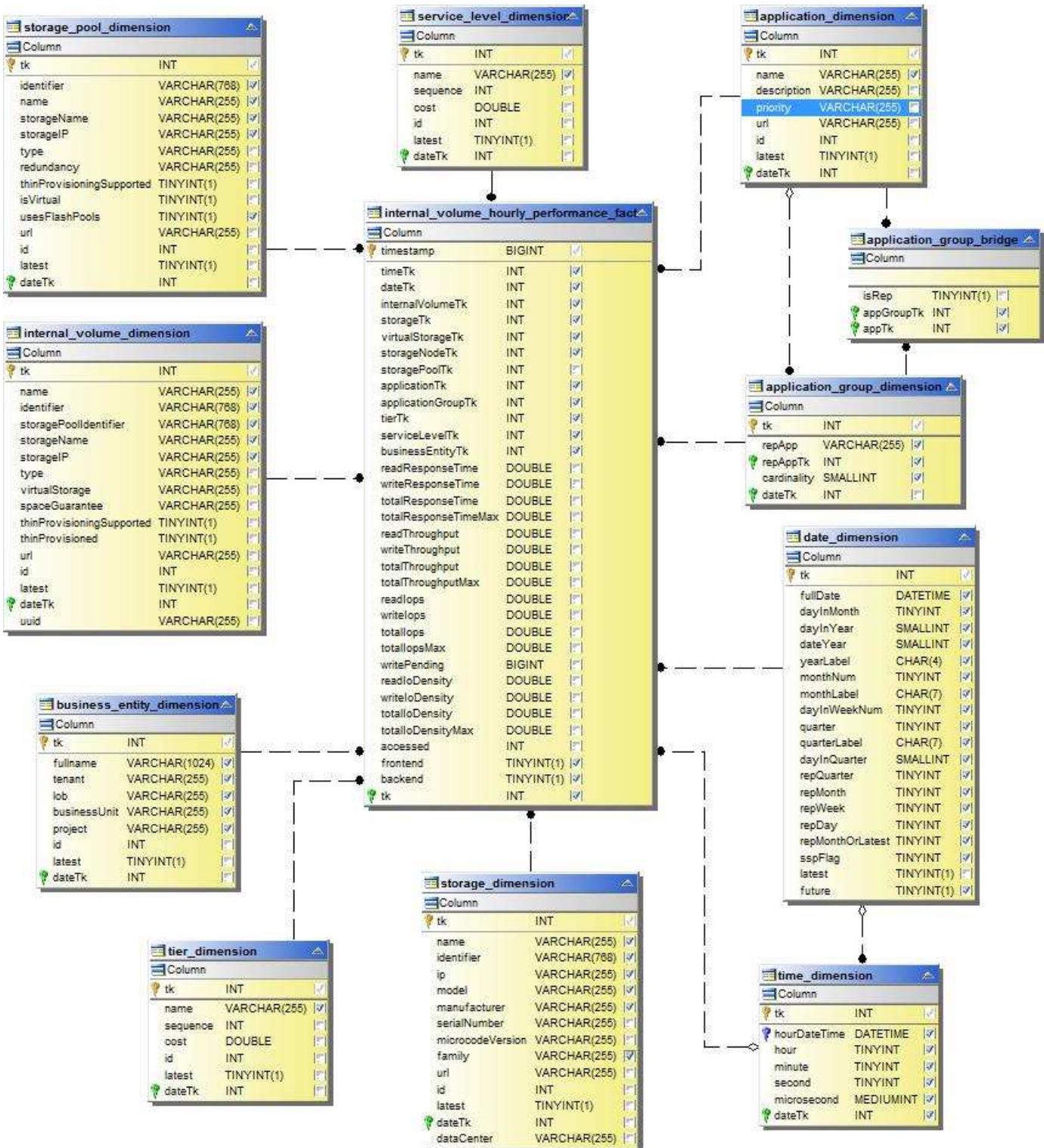
Application Volume Hourly Performance



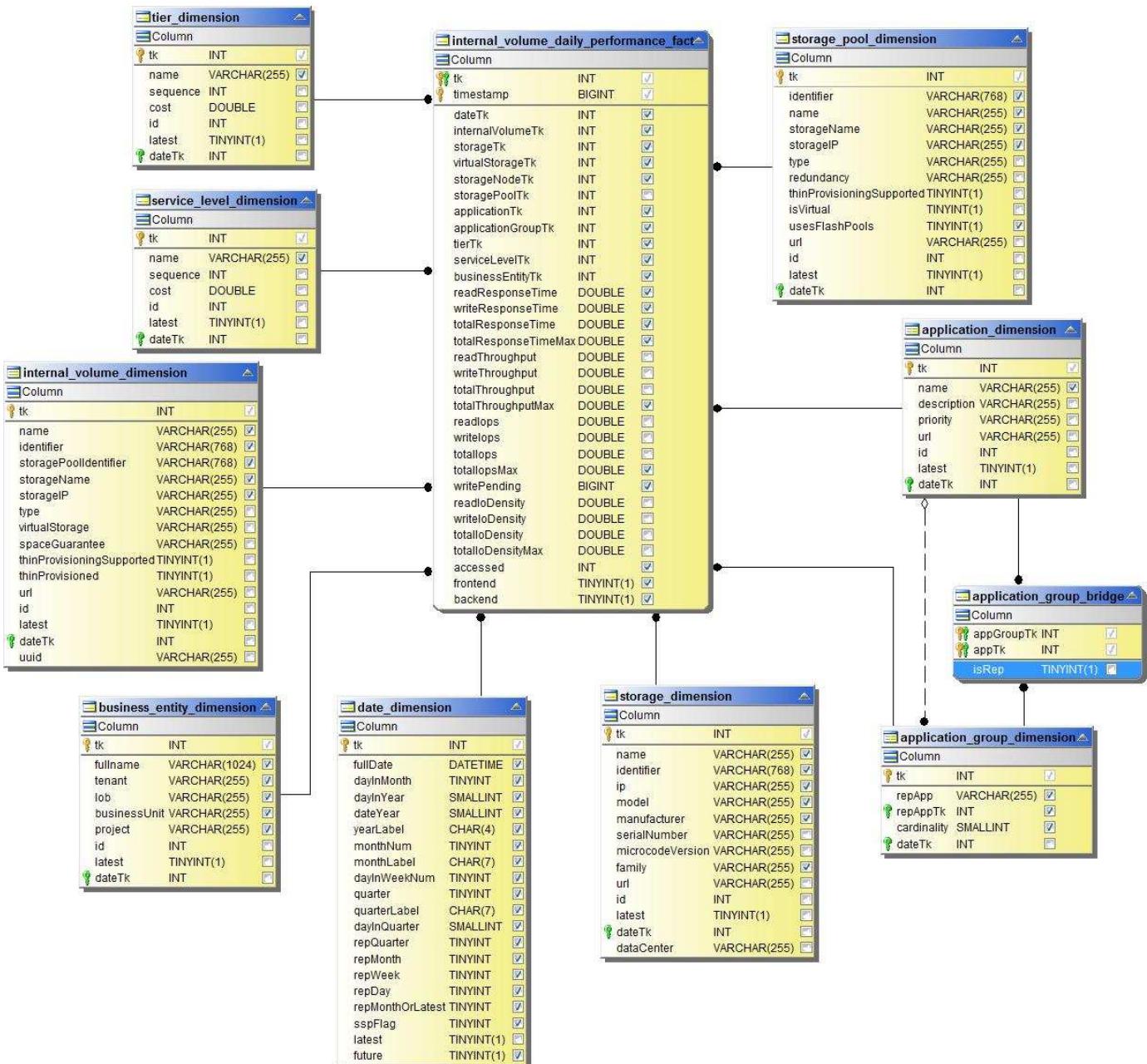
Internal Volume Hourly Performance



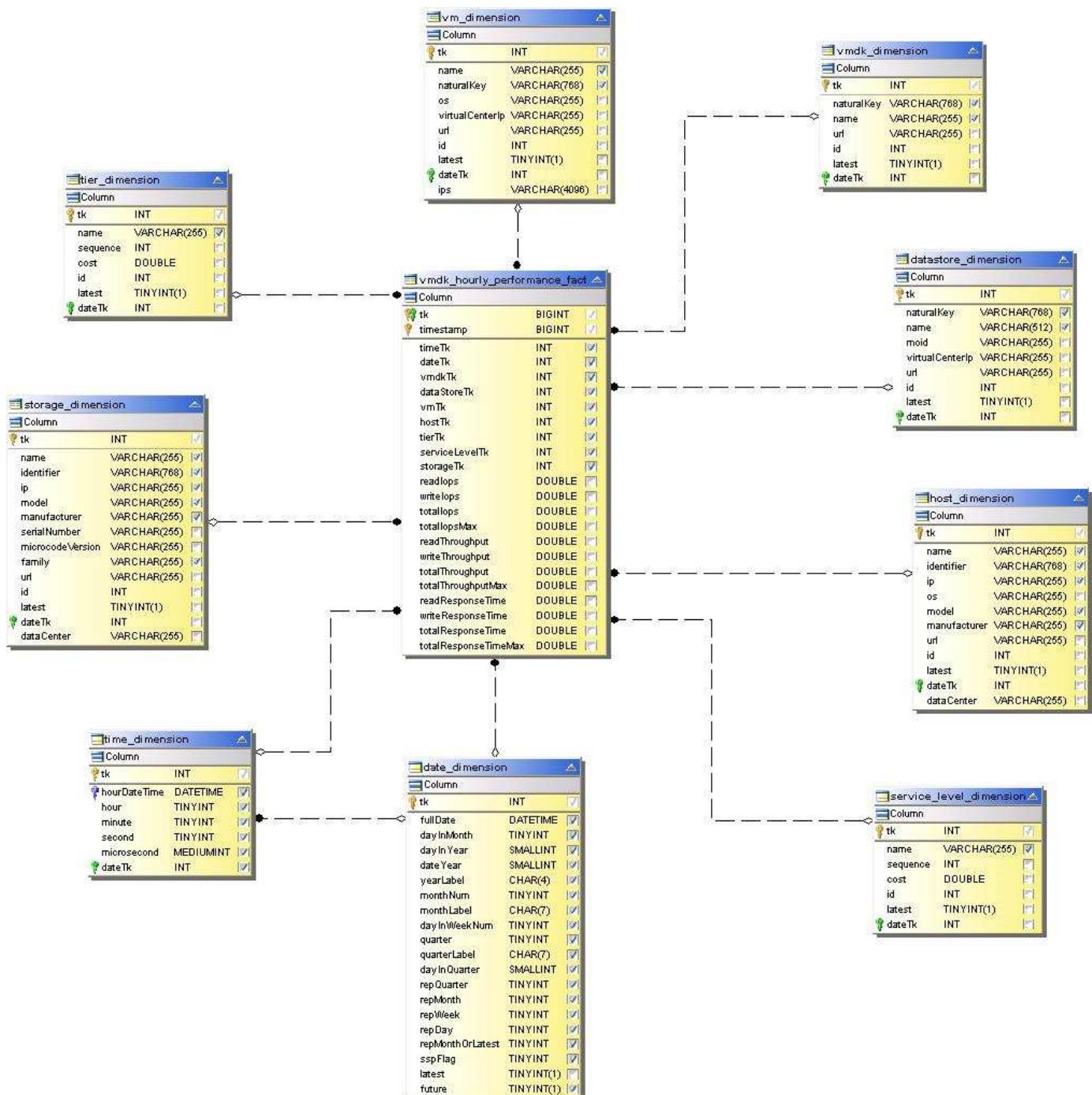
Internal Volume Hourly Performance



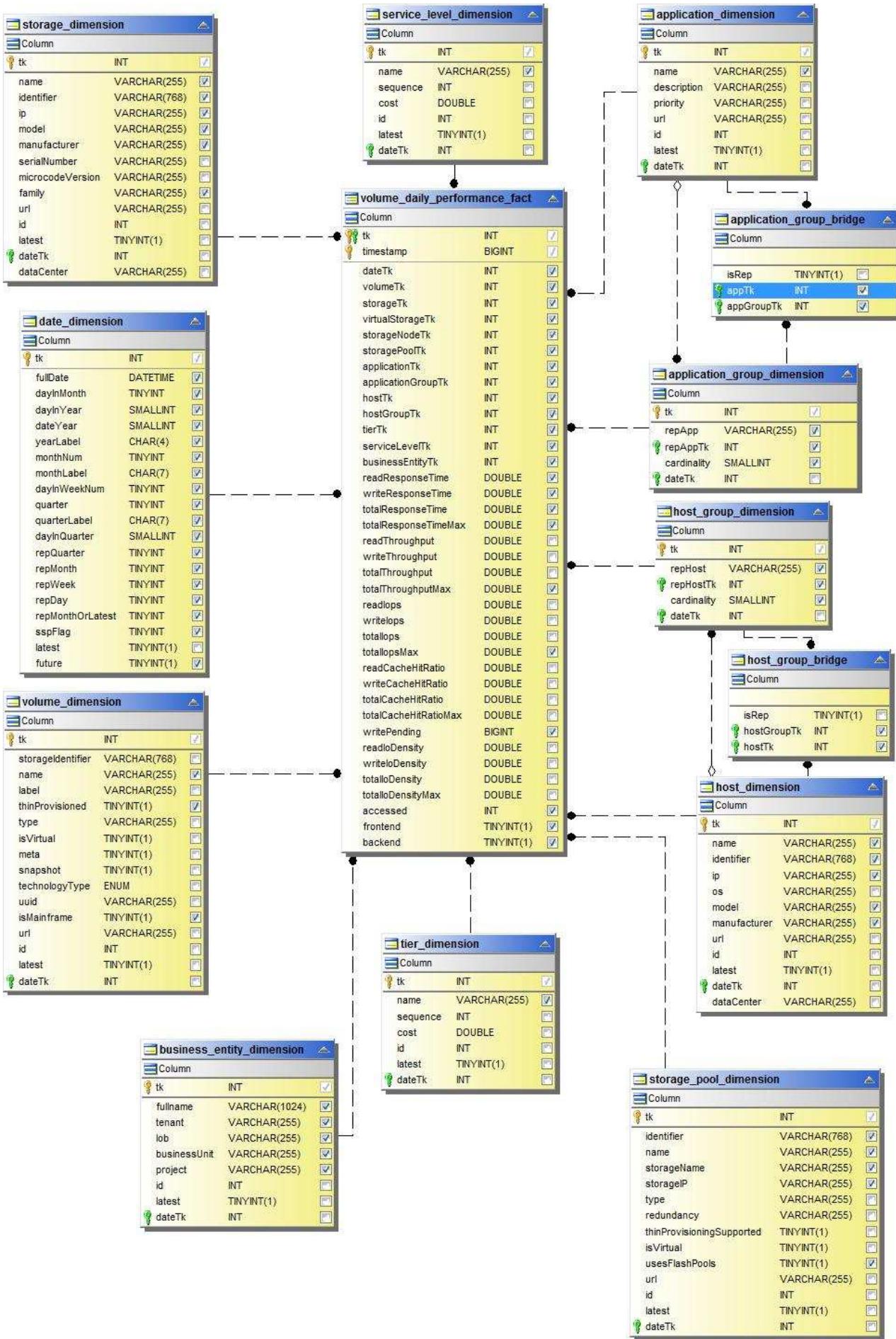
Internal Volume Daily Performance



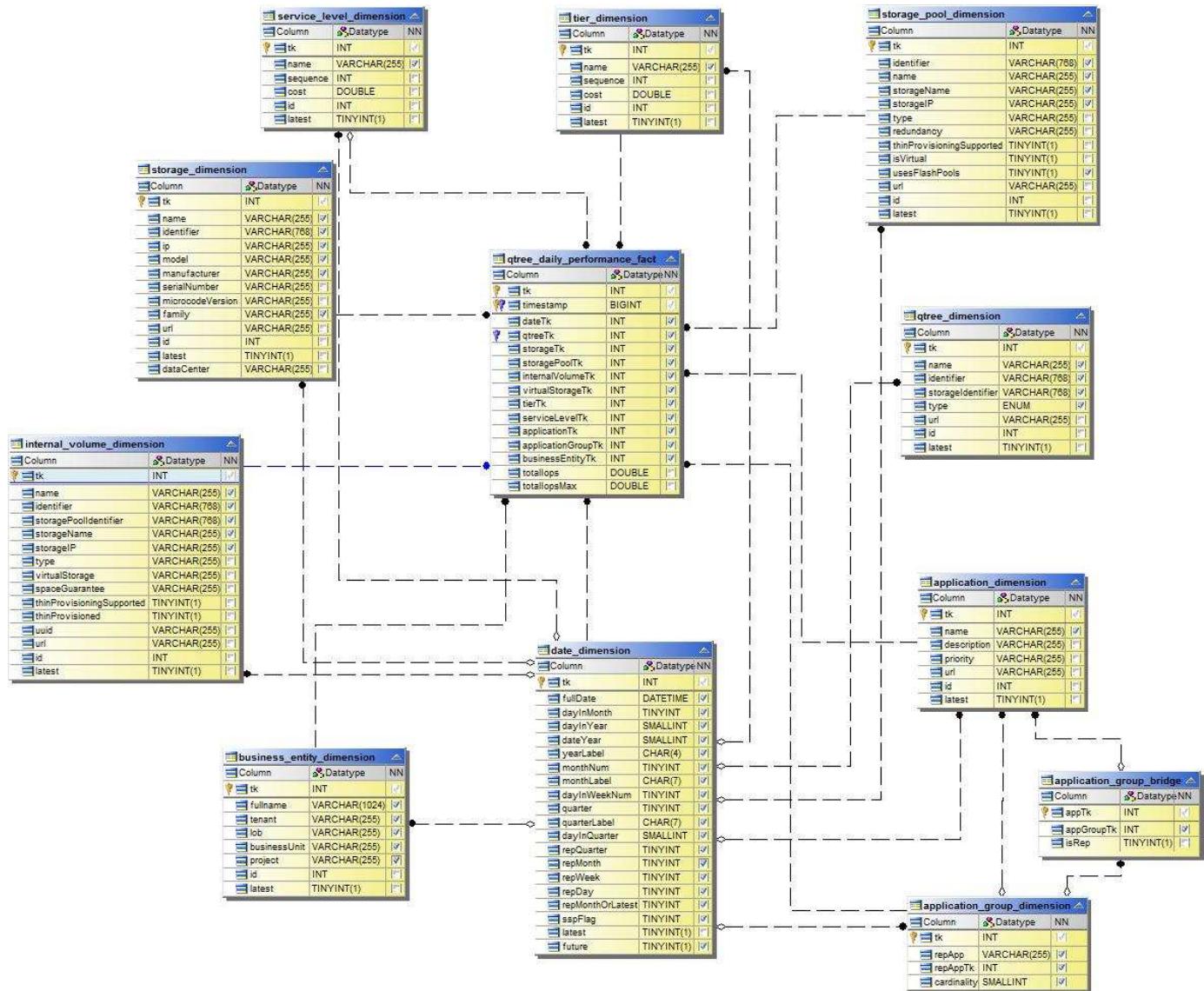
Volume Hourly Performance



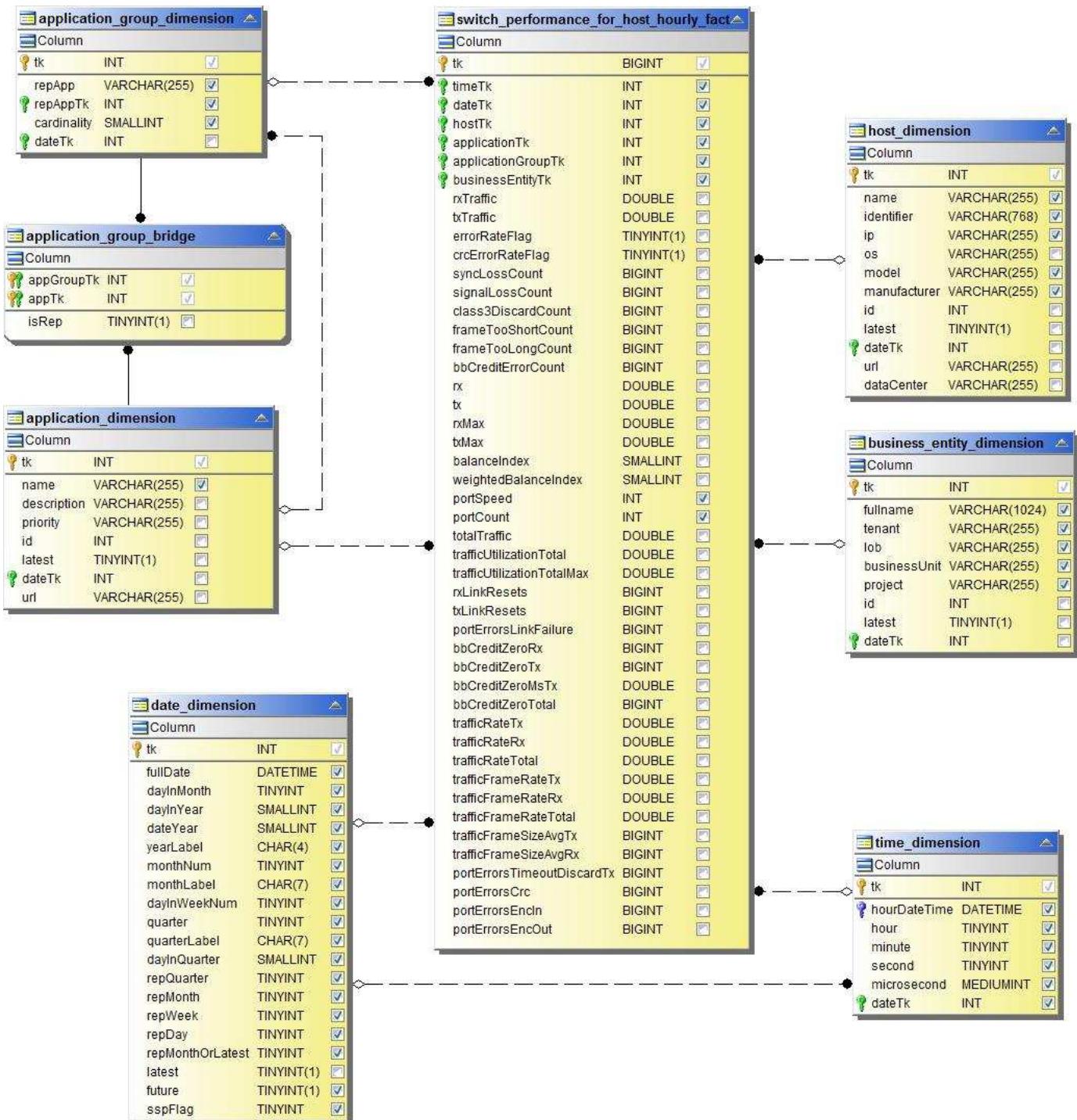
Volume Daily Performance



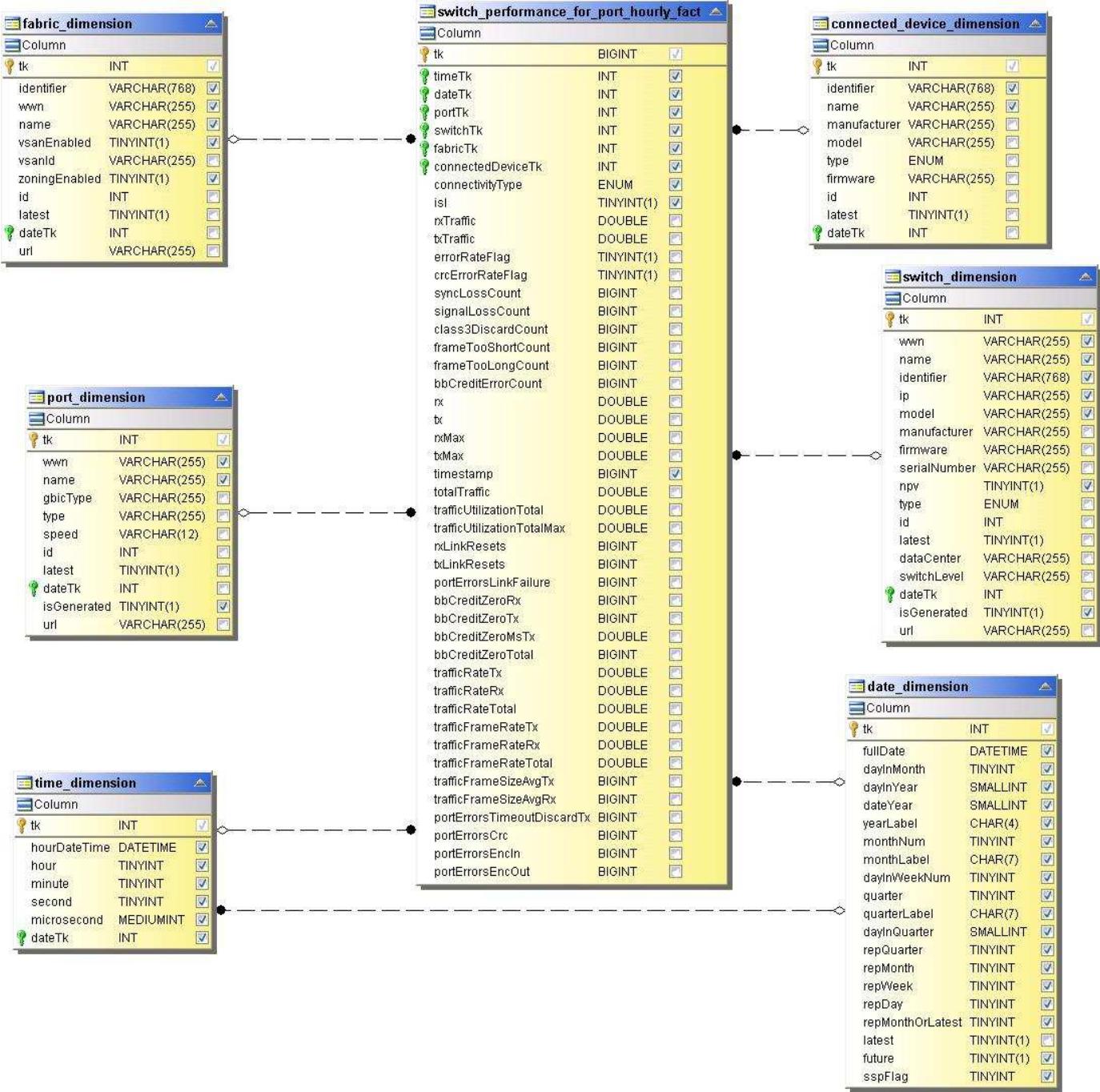
Qtree Daily Performance



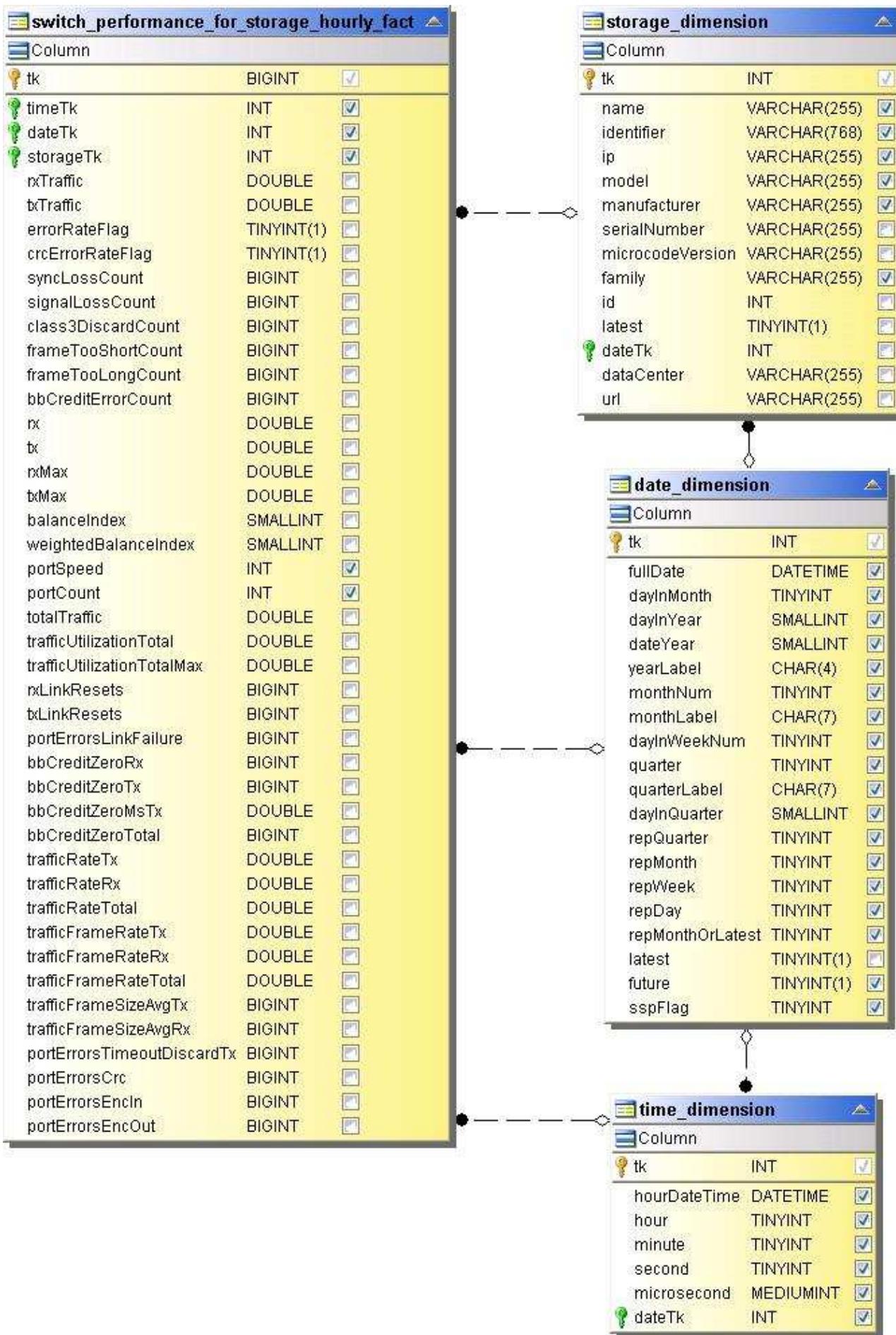
Switch Hourly Performance for Host



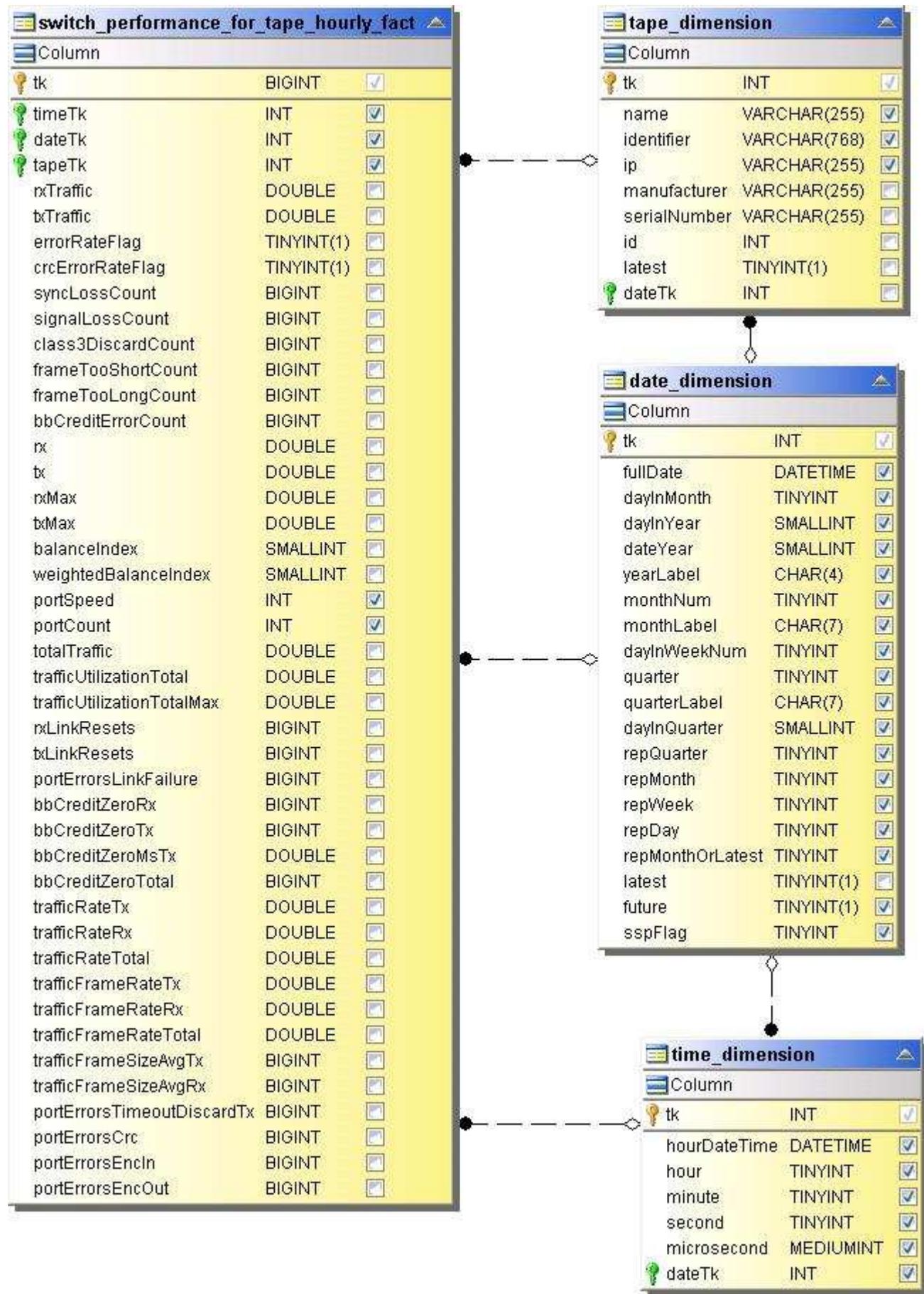
Switch Hourly Performance for Port



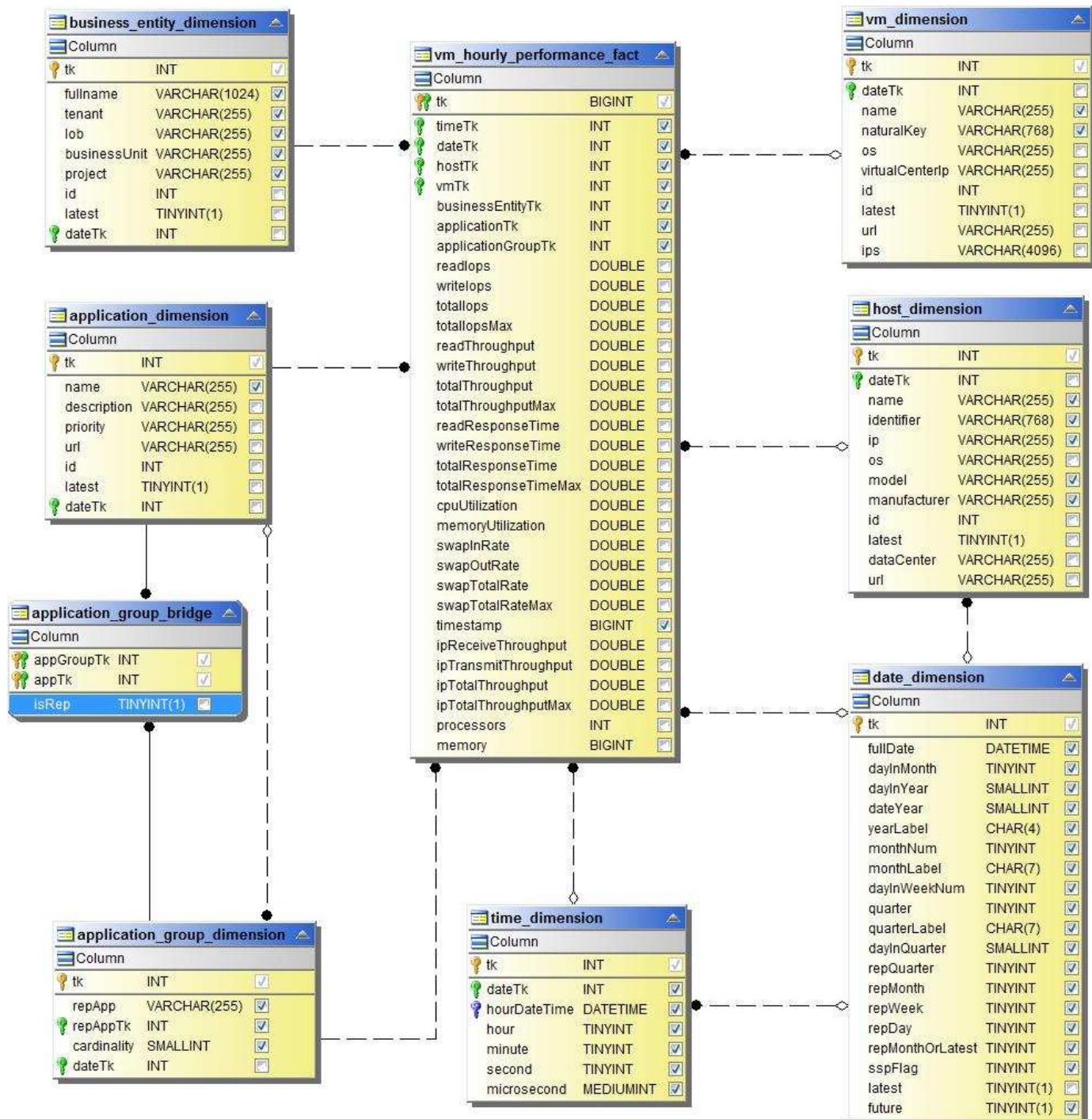
Switch Hourly Performance for Storage



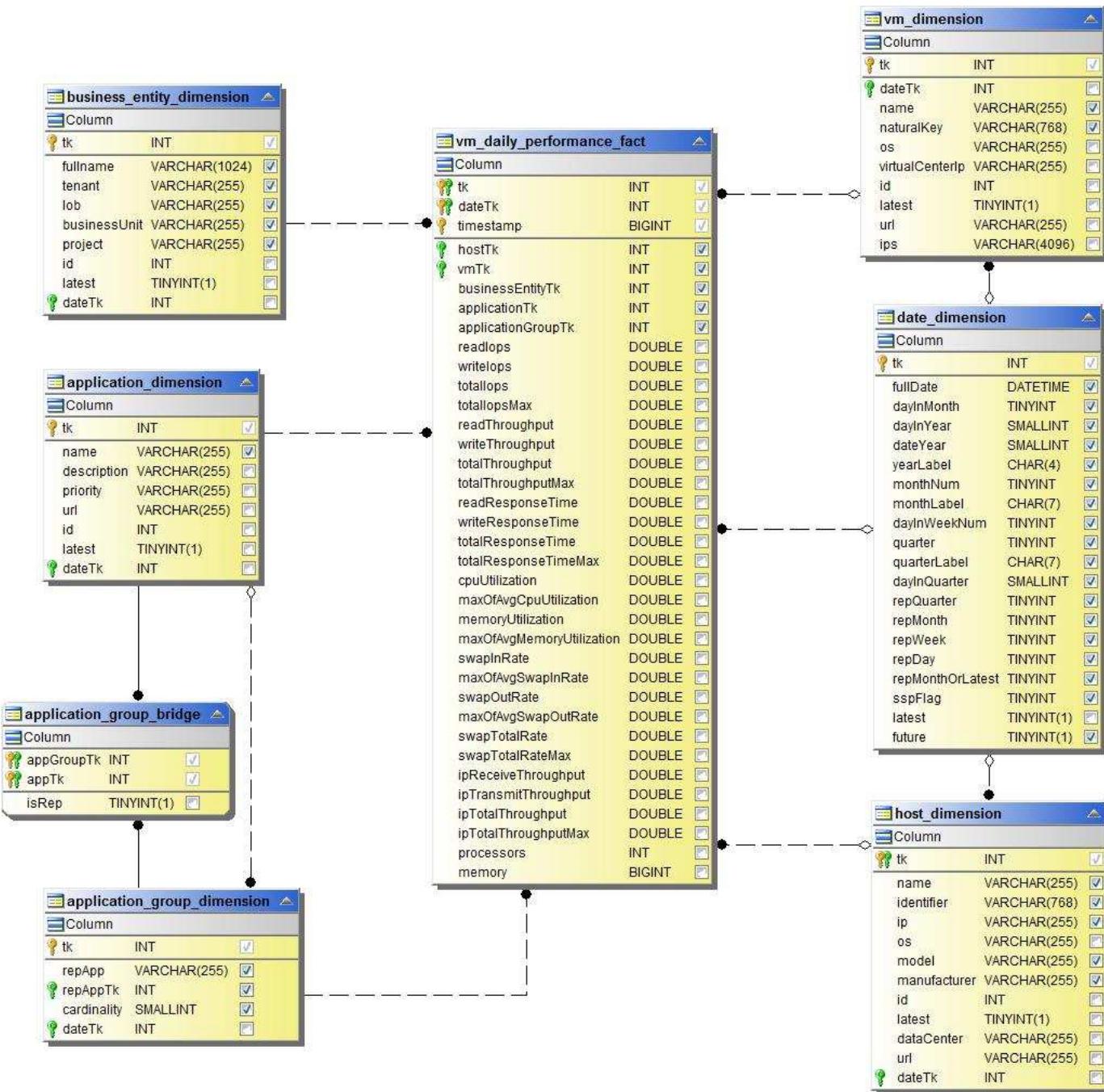
Switch Hourly Performance for Tape



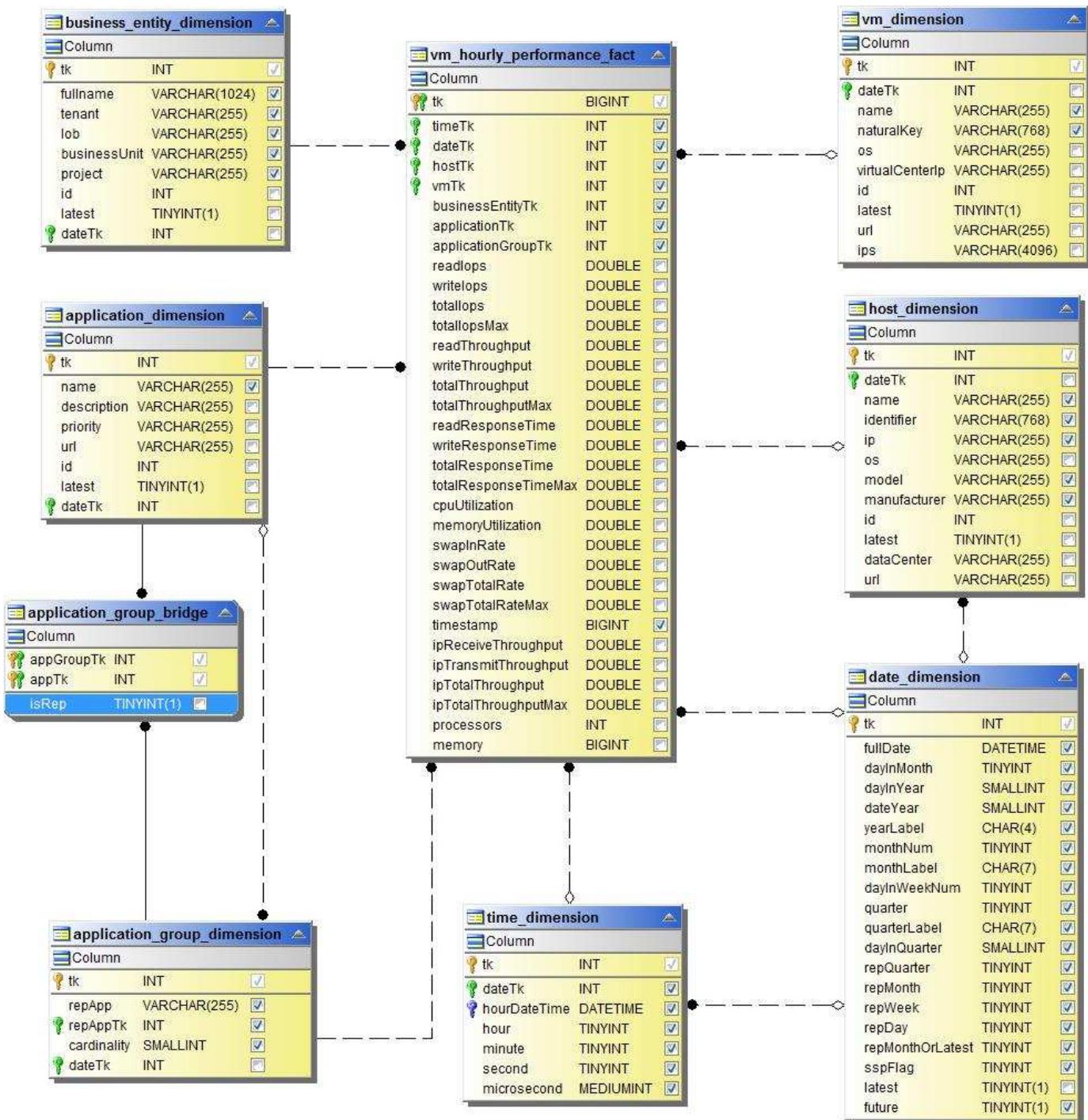
VM Performance



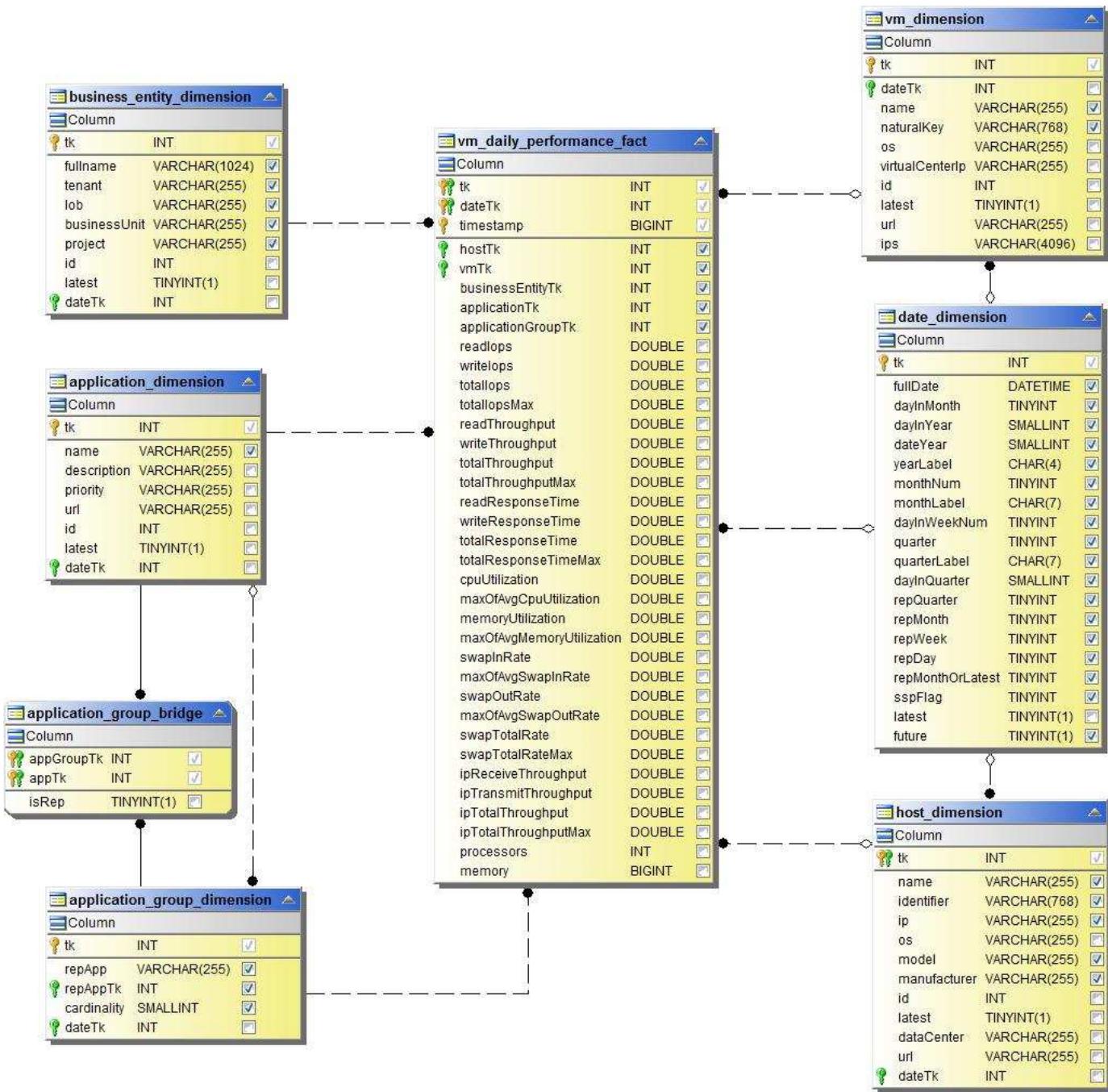
VM Daily Performance for Host



VM Hourly Performance for Host



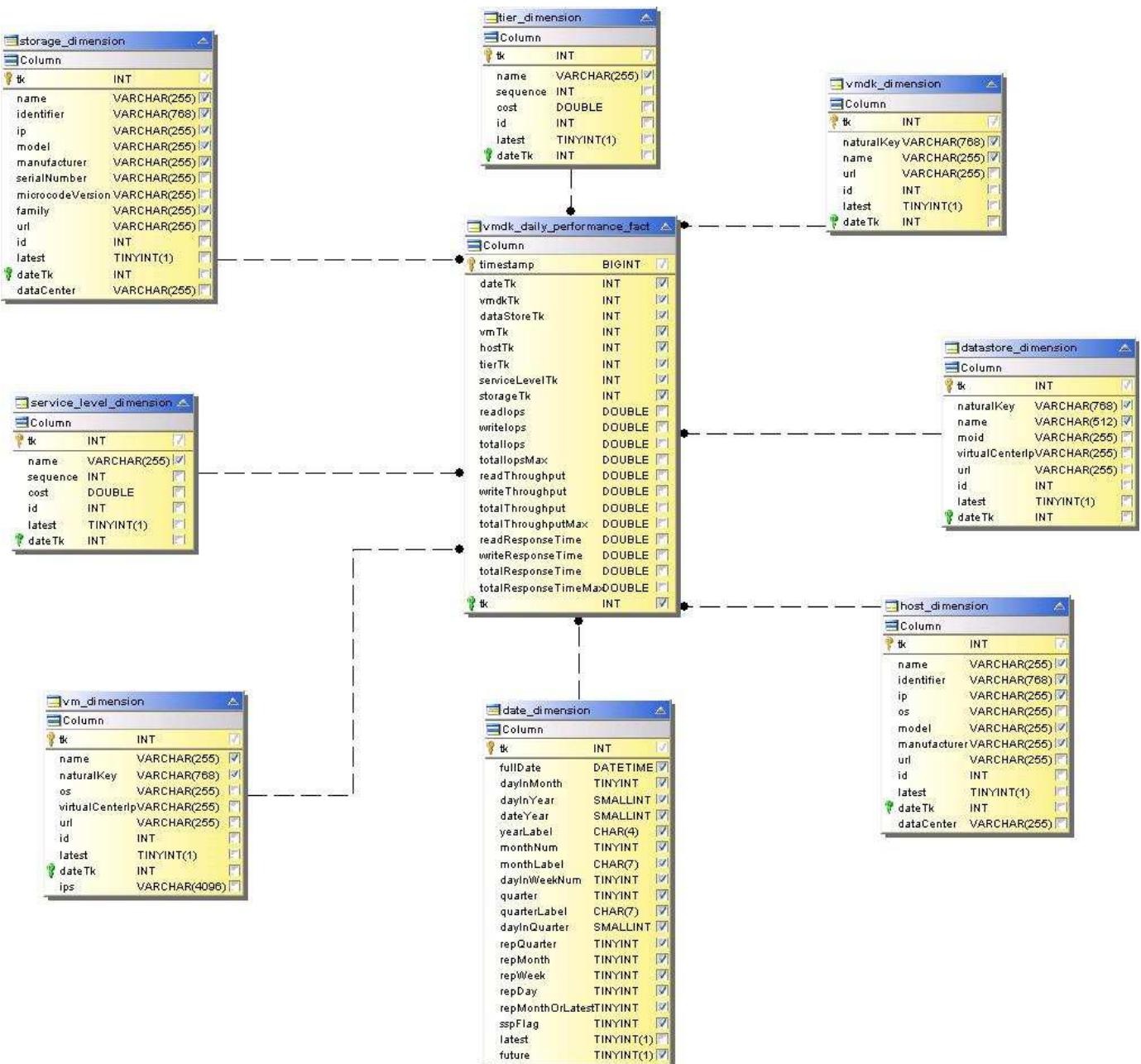
VM Daily Performance for Host



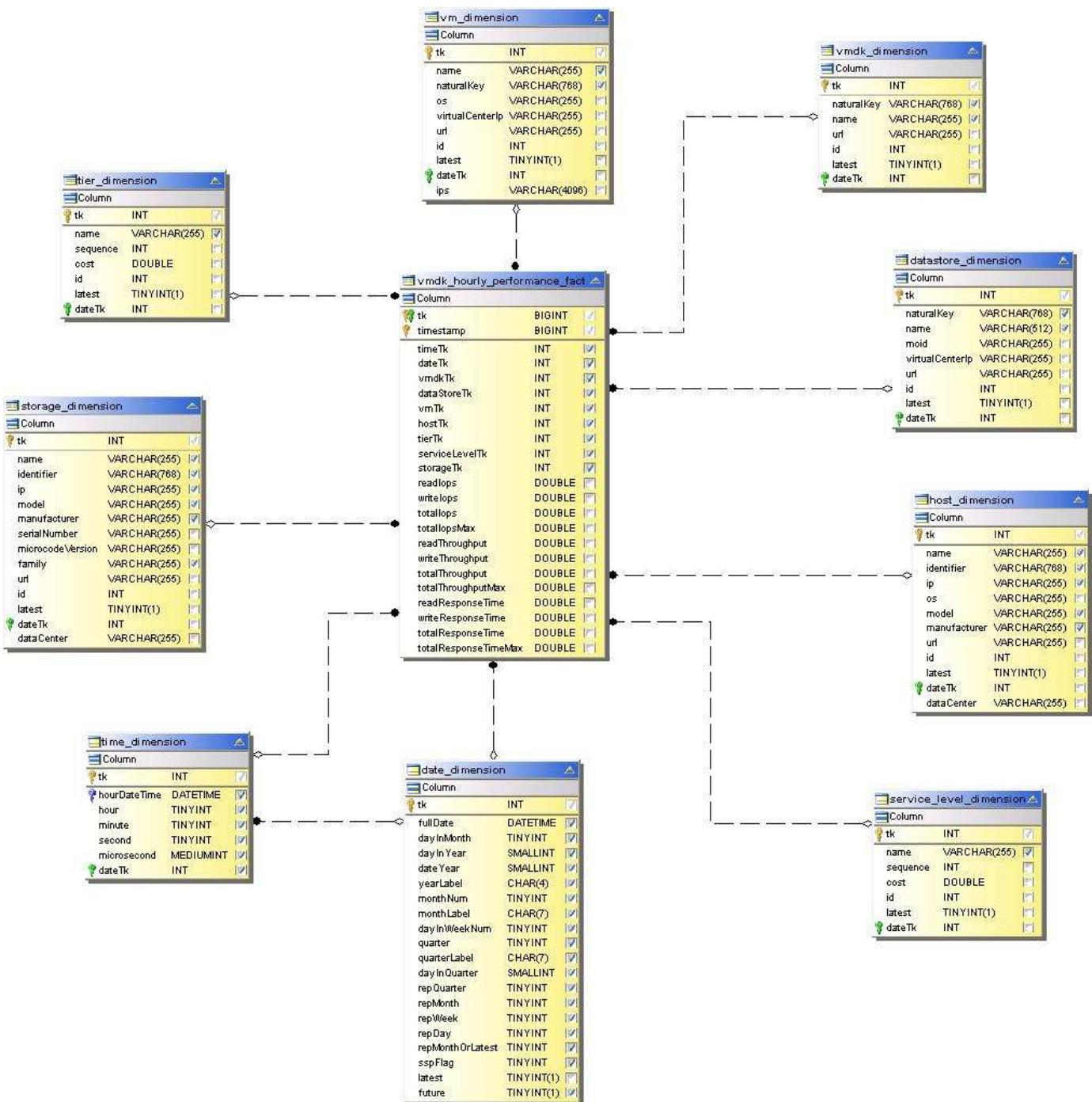
VM Hourly Performance for Host



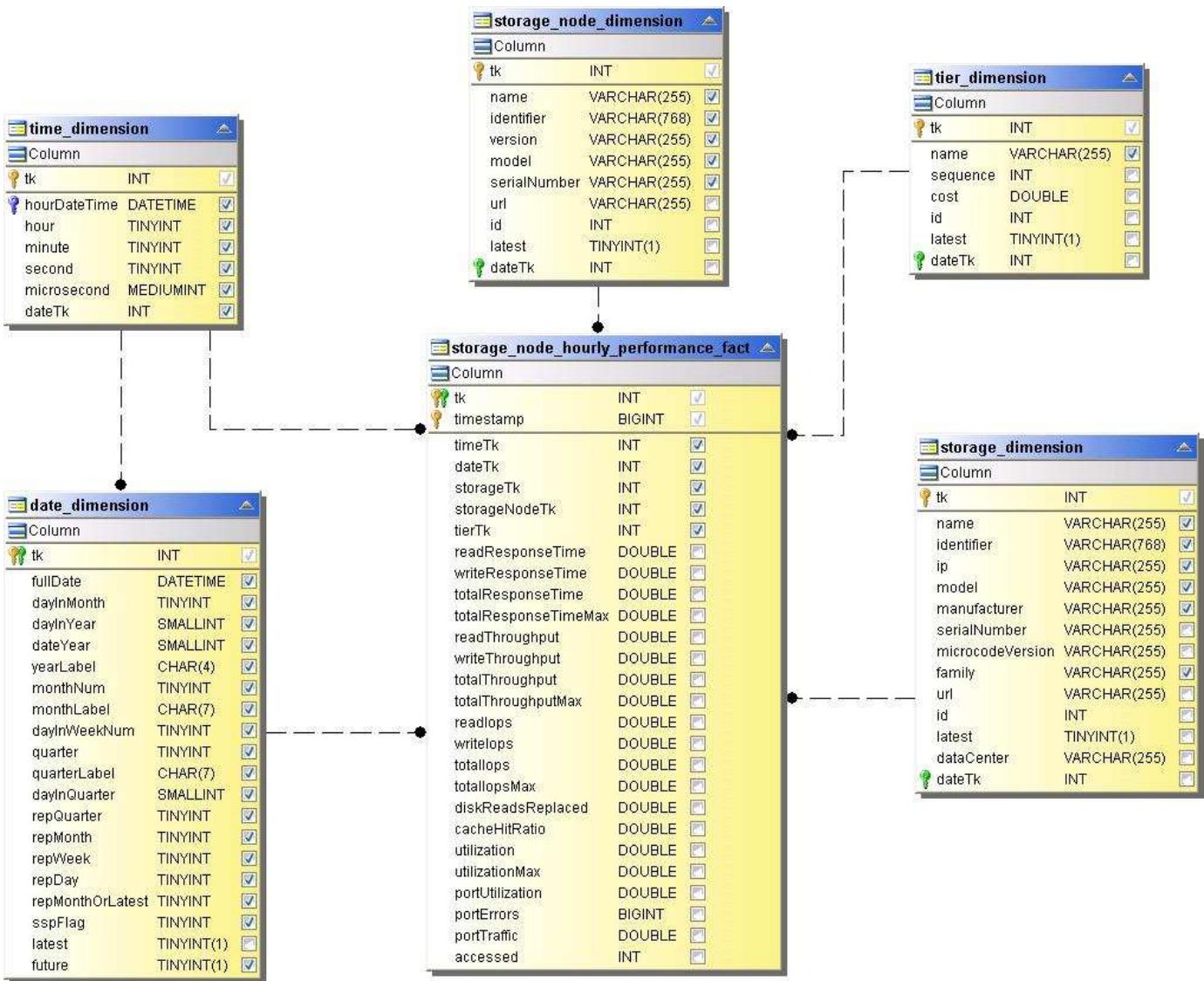
VMDK Daily Performance



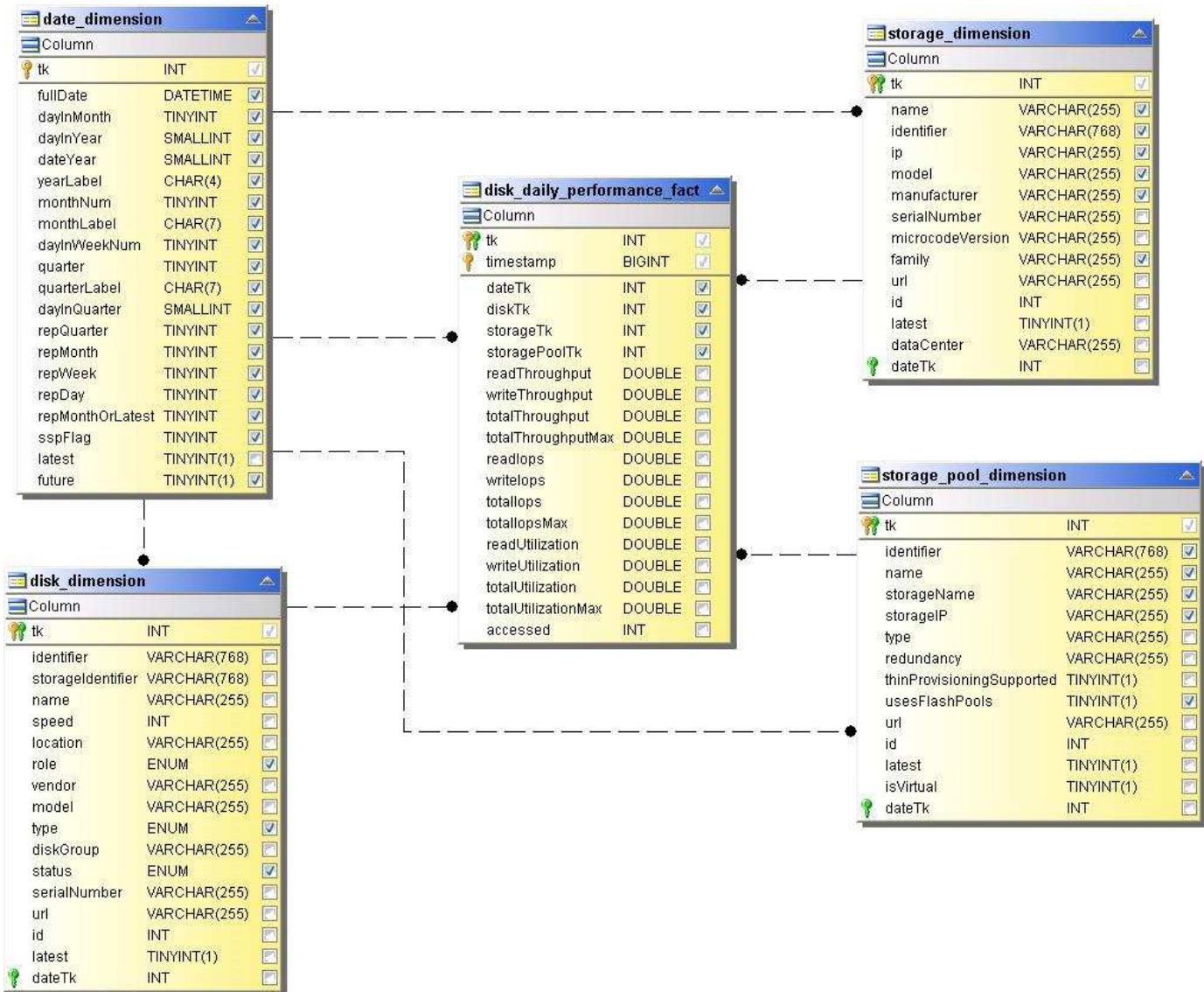
VMDK Hourly Performance



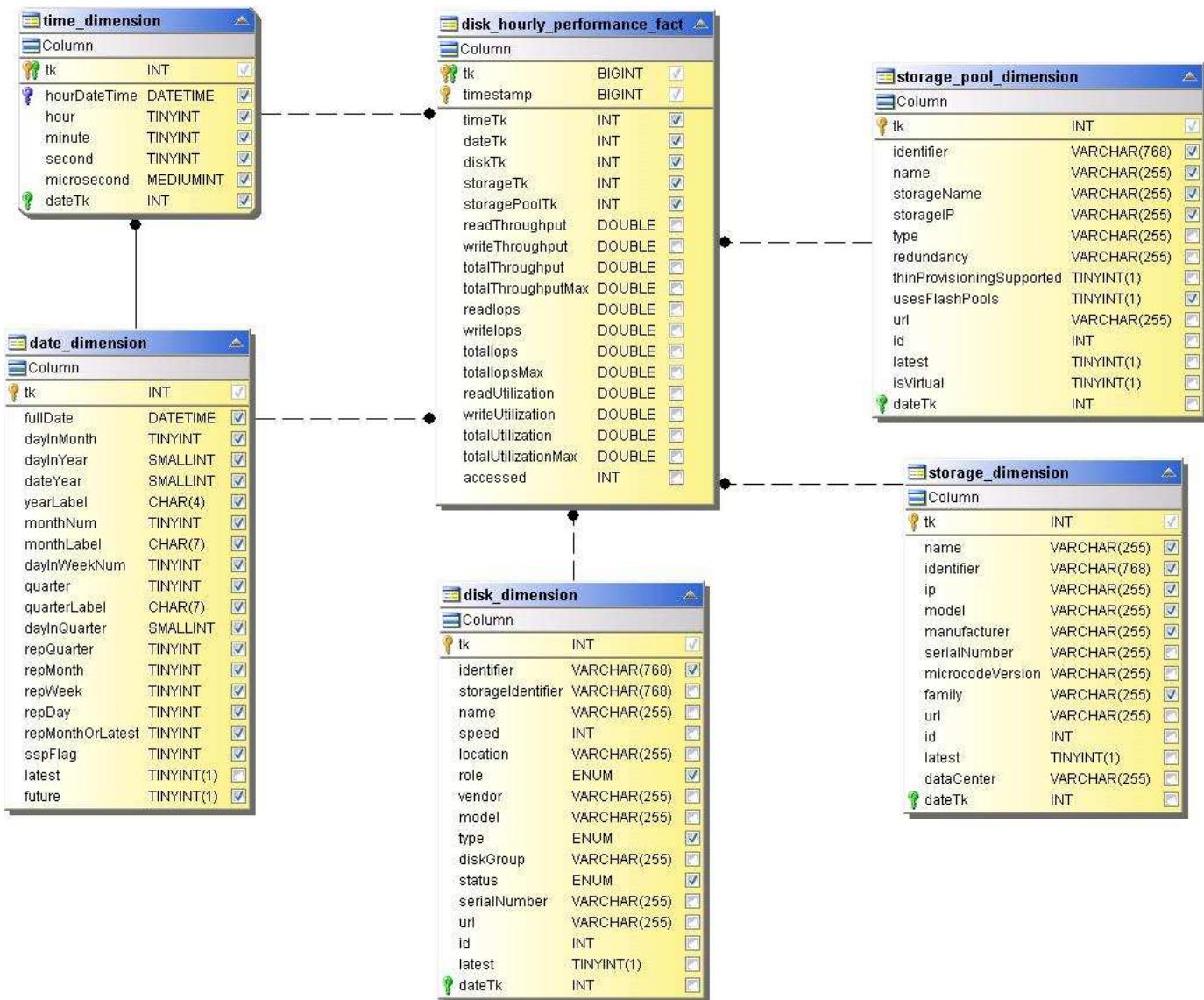
Storage Node Hourly Performance



Disk Daily Performance



Disk Hourly Performance



Cloud Insights Schemas for Reporting

These schema tables and diagrams are provided here as a reference for Cloud Insights Reporting.

[Schema Tables](#) in .PDF format. Click the link to open, or right-click and choose *Save as...* to download.

Schema Diagrams



The Reporting feature is available in Cloud Insights [Premium Edition](#).

Copyright Information

Copyright © 2021 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means-graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system-without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

Trademark Information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.