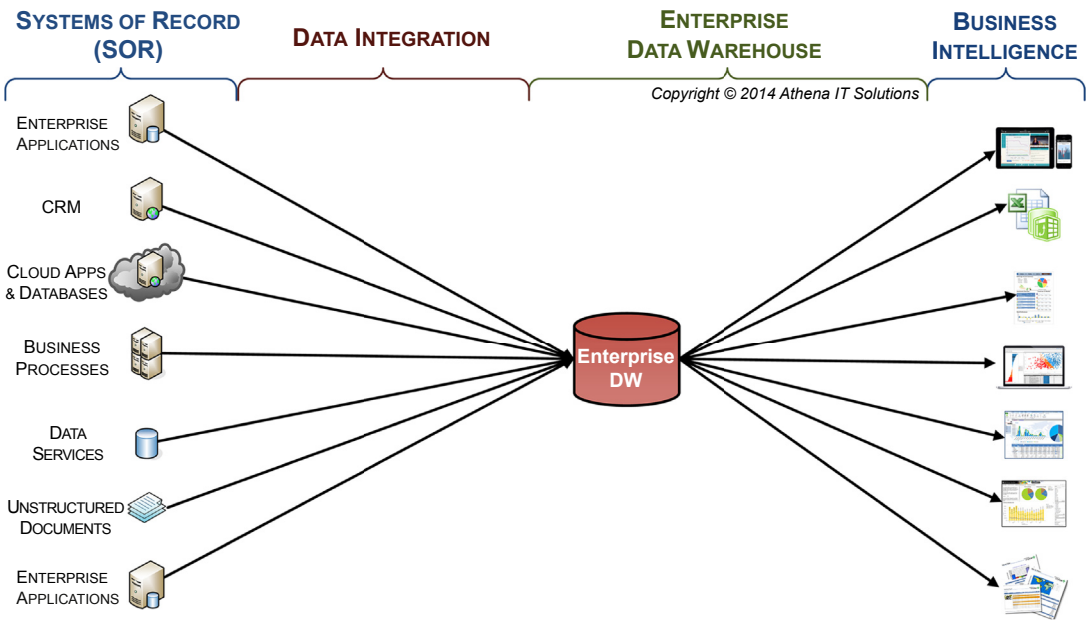

FIGURE 6.1

Enterprise data warehouse (EDW).

Designing, building, and implementing CDWs took a long time—months and maybe years. Some of the reasons for the lengthy time to deliver a CDW were:

- IT's use of the waterfall project methodology
- Everything had to be manually coded
- Much of the database, access, and networking standards that we take for granted today had not yet evolved, requiring the use of specific propriety APIs

The biggest reason why CDW projects took so long was IT's misguided attempt to gather all the requirements and implement an entire CDW in one "big bang" project. With an enterprise-wide scope, the project was destined to consume many resources, be late, and exceed the budget. An even worse outcome was that while the lengthy project was under way and after business requirements were documented, the business changed—invalidating many of the requirements.

Being an early adopter always carries some risk, and this was no exception. In addition to missing deadlines and running over budgets, CDW projects tended to under-deliver on expectations. And that was for the projects that actually were completed; many simply fizzled out. It is worth noting that many other big corporate projects in the 1990s encountered similar fates. There were scattered success stories, but, in general, the underwhelming results gave CDWs a bad reputation. It was inevitable, therefore, that there would be a backlash.

THE DATA MART

The next era was the rise of data marts. Data marts promised to be quicker and cheaper to build, and provided many more benefits—including the benefit of actually being able to finish building them! The data mart was

primarily a backlash to the big, cumbersome CDW projects, with the key difference being that its scope was limited to a single business group rather than the entire enterprise. Of course, that shortcut did speed things up, but at the expense of obtaining agreement on consistent data definitions, thereby guaranteeing data silos.

Although the early-adopter technology companies and Bill Inmon both advocated that data marts should be fed from the EDW, the rest of the industry was pushing data marts based on a different architecture where they were fed directly from source systems. Instead, data mart vendors and pundits stated there was no real architectural difference between a DW and data mart. As Figure 6.2 shows, the data mart, just like the DW, sourced data from SORs, used a waterfall project methodology and had to use quite a bit of manually-created custom code. The pundits' sales pitches said that the data mart was just like a DW – using the same tools, data models and design principles – but simpler and faster to build.

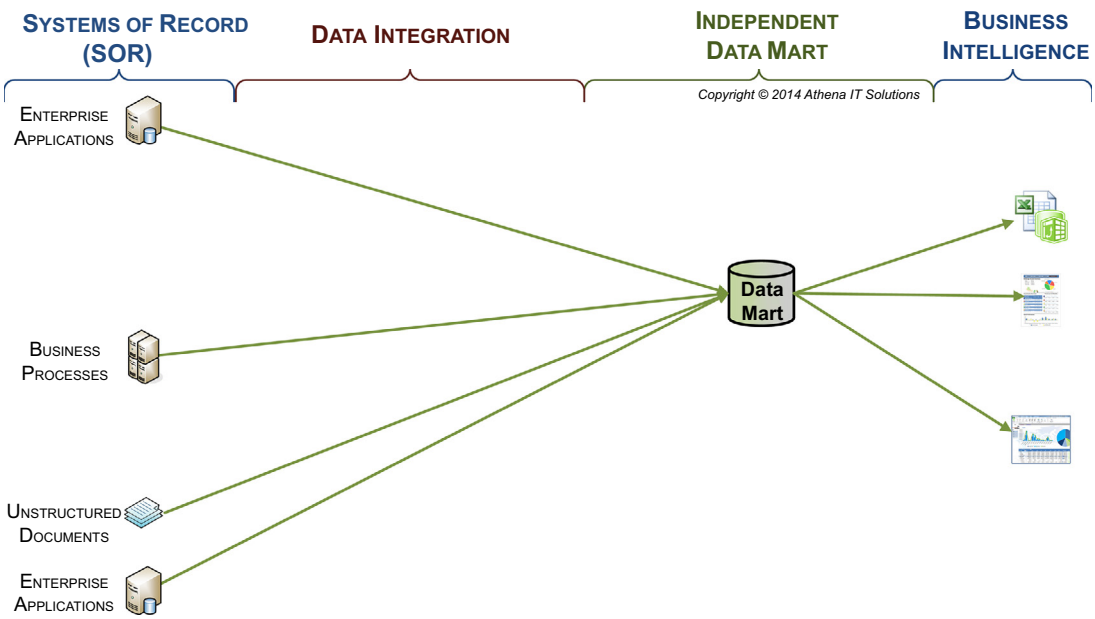

FIGURE 6.2

Data mart.

By the time data mart projects were starting to be widely built, two trends were also occurring: the emergence of BI tools and the use of relational databases to build new source systems. These trends resulted in a much shorter time to build a data mart than a CDW, aided, of course, by the fact that data marts have a more limited scope than CDWs.

Rather than being part of an overall enterprise data architecture, data marts were the architecture. You could build them faster and cheaper because you did not take the whole rest of the enterprise into account. This meant you could take shortcuts and avoid the most difficult part of DW: DI and reaching agreement on data definitions and metrics across business groups. As a result, data marts were built for the parochial interests of their sponsors rather than for the enterprise.

Companies sprouted multiple data marts, each built to accommodate only their sponsor with their own data definitions, data transformation, reporting, and technical platform. We moved from trying to