| CASE STUDY
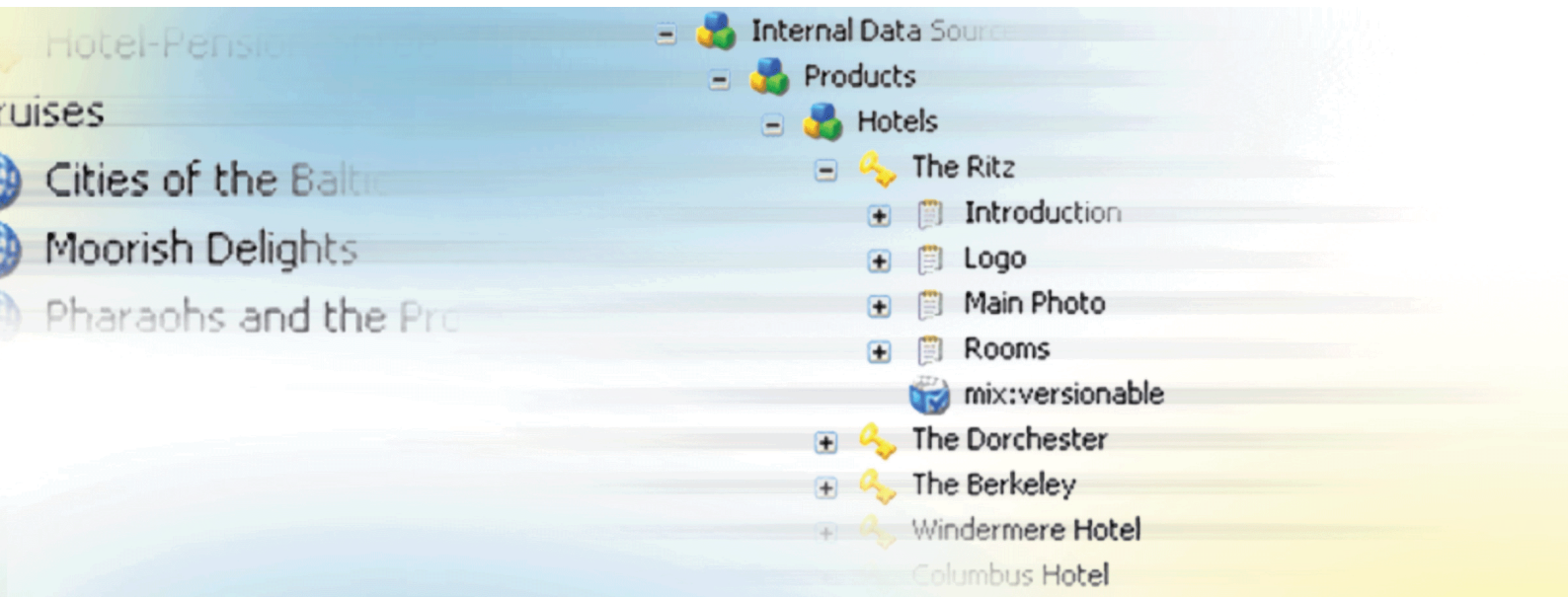
TUI Travel – A Scalable Content Platform

BACKGROUND

TUI Travel is the world's leading travel leisure company, serving more than 30 million customers in over 180 countries. Its 200+ brands generated over £13bn in revenue during 2010, and TUI Travel is the only travel leisure company listed on the FTSE 100 index.

TUI operates brands such as Thomson, LateRooms, FirstChoice, Marmara, Hotelopia, Arke and Mostravel, and owns hotels, resorts and a large fleet of aircraft.

Many specialist content providers – such as LonelyPlanet, GIATA, and Weather Underground – supply content and media to the travel industry, due to the need to maintain fresh and up-to-date media as hotels change hands, or are redecorated, or new attractions or resorts become available.

Following a period of expansion and acquisitions, TUI found that many of its websites were sourcing the same content (such as images of hotels or destinations, videos, geo-data and weather data) from the same suppliers, each website paying not only for licenses to use the content but also for the integration effort between their systems and the content sources.

Each website or market thus incurred complexity, cost and effort in order to integrate with content sources, store the content (often many Terabytes in size), and operate the corresponding systems. Each website was effectively "re-inventing the wheel", solving problems which other TUI group businesses had already solved.

THE CHALLENGE

TUI Travel needed to be able to arrange group content deals with content providers in order to reduce content licensing costs, and therefore needed a central repository in which to manage and store the shared digital assets. TUI also needed to reduce time-to-market for new web-based products and services, and to avoid repeated system integration activities between content systems, whilst taking advantage of shared business data.

It became obvious early on in the project that the solution to TUI's business problem could not be met with off-the-shelf software. For instance, a Web Content Management system would not be suitable due to the need to be presentation-independent. A Digital Asset Management (DAM) package would not provide the required integration with business data, nor the content packaging required. An ecommerce-type Product Information Management (PIM) solution would emphasise too strongly a product-based approach, where a focus on content and entities was required. Other technologies, such as the SharePoint Business Data Catalogue (BDC) were too specialised.
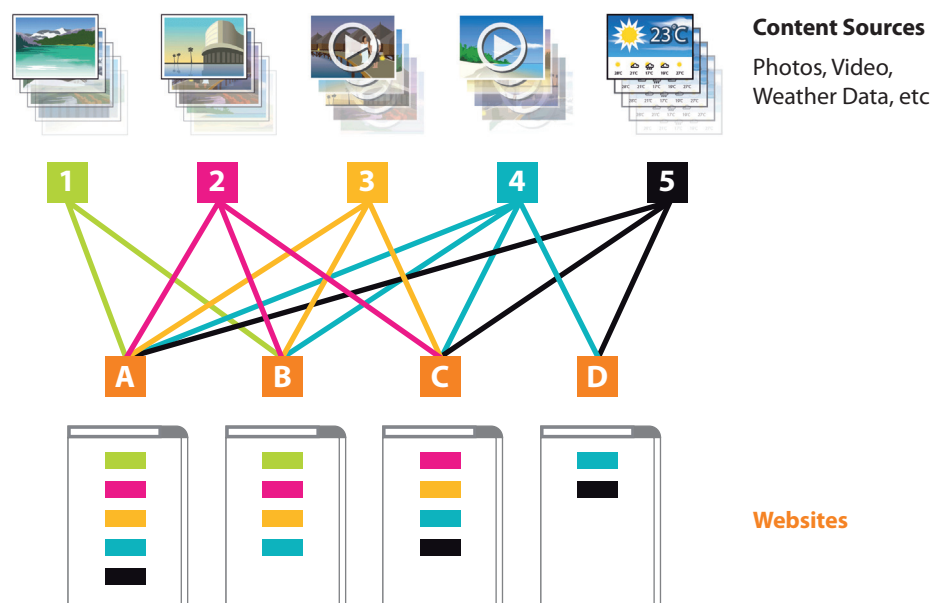

Figure 1 - Multiple, repeated integrations and license deals

THE SOLUTION

TUI Travel approached Priocept in 2008 for help with designing and building a technology system, the TUI Content Platform, to meet their needs.

The system was to act somewhat like a "digital concierge" – providing information and data in the user's preferred language, rapidly and efficiently.

Instead of each website having to connect to each content source directly, the Content Platform aggregates the content, making it available via a standard API to each website, thereby simplifying integration, reducing license costs and adding value in areas such as content packaging, access control, high-performance caching, and integration with Master Data.

Priocept also ran a vendor selection process for the hosting provider, eventually selecting Infotec, a hosting company based in Hamburg, Germany. Infotec's offering, using VMware ESX 4, Red Hat Enterprise Linux, and Squid Cache, stood out as highly configurable whilst offering excellent performance characteristics.

In consultation with the E-Commerce Architecture team at TUI, Priocept designed a system which provided:

- A services-based content "hub"
- High-performance, scalable architecture for content distribution
- Master Data (MDM) integration
- Intelligent caching and statistics
- Advanced access control and security
- Multi-lingual, multi-media integration interface (API)
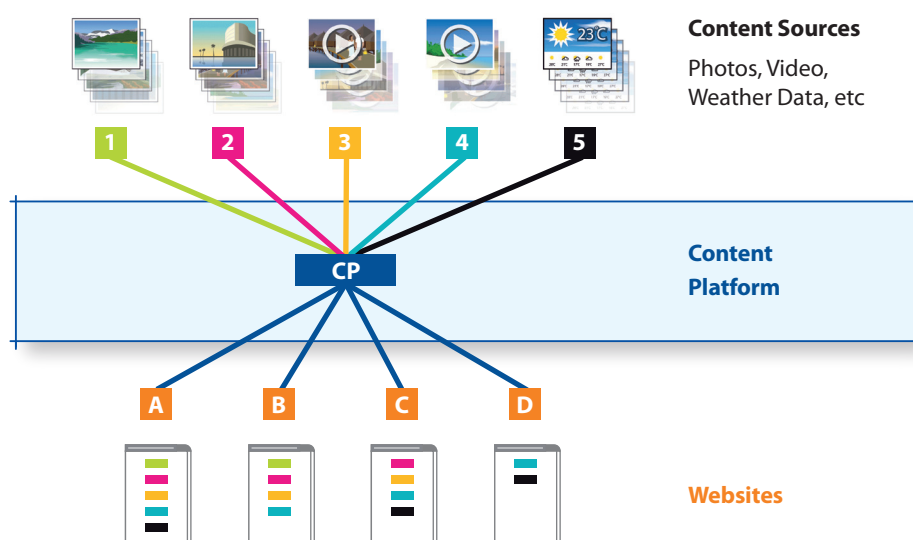- Content manipulation and packaging

Figure 2 - The Content Platform as the "hub", simplifying integration and content packaging

TECHNOLOGY

The resulting solution was based on an industry standard set of technologies:

- Java 6
- Spring and CXF
- Tomcat
- Jackrabbit (JCR)
- MySQL 5.1
- Squid Cache
- Red Hat Enterprise Linux
- VMware ESX 4

Java content repository (JCR)

JCR is an open standard, originally developed by Day Software (now part of Adobe Digital Enterprise Platform), and designed specifically to simplify the design and build of content-focused systems by providing a common language and set of techniques for interacting with repositories of content.

Two products stood out as candidates for use in the TUI Content Platform: Day CRX (a licensed product from Day Software, acquired in 2010 by Adobe), and Apache Jackrabbit (an open-source technology). In fact, Day CRX is built on top of Apache Jackrabbit, and extends it with various additional features, such as a SAN-based TAR persistence manager for high-performance and database-free

repository access, and CRX Explorer, an advanced repository browser which allows administrators to configure, search and manage JCR repositories.

Priocept undertook extensive prototyping with both CRX and Jackrabbit, including extensive performance tests using different repository configurations. As a standalone JCR technology, the main advantages of CRX over Jackrabbit are the extra repository management features, performance enhancements, and commercial support. However, in this case, Jackrabbit with MySQL as the data store was considered sufficient for the project.

Priocept went on to design and develop a content repository browser for Jackrabbit JCR repositories, similar to the CRX Explorer from Day, called Jackrabbit Explorer.

BENEFITS

The combination of the various technologies provided a powerful platform on which to build a scalable and dynamic system. JCR/Jackrabbit & MySQL provided the core advanced data store, including querying, access control, versioning, replication, and clustering.

Java/Spring/CXF enabled implementation of flexible, well-versioned web services backed by domain-specific logic (such as multiple geographic views e.g. "are the Canary islands in Europe or Africa?") and server-side mashups (image galleries, image resizing and watermarking, content fallback, etc).

Using Squid Cache allowed transparent caching of both local and remote content, and optimised content expiry for different scenarios, such as weather data - this helped the system scale to about 1 Terabyte of content and about 200k hotels and 70k locations.

The result – TUI Content Platform – is a scalable, configurable and flexible services-based intelligent content hub, integrating with the TUI Group Master Data reference data solution. The system loads and categorises content automatically, and provides a range of options for integrators from the 200+ TUI Group websites, enabling them to make detailed requests for relevant content, which is then served rapidly through the caching engine.

Any of the 200+ TUI consumer-facing websites can now connect into the Content Platform using a simple web services API and retrieve rich, multi-lingual content in many different formats for thousands of hotels and destinations around the world, without the complexity of integrating with each content source separately, resulting in significant time and cost savings.

" The quality of Priocept's technical concepts as well as implementation is first-rate. Priocept staff have consistently provided valuable feedback and suggestions for improvement during concept and implementation phases while staying absolutely focussed on delivering the solution in time, quality, and budget. "

Dr. Karl Geiger
Head of Web Development
TUI Travel
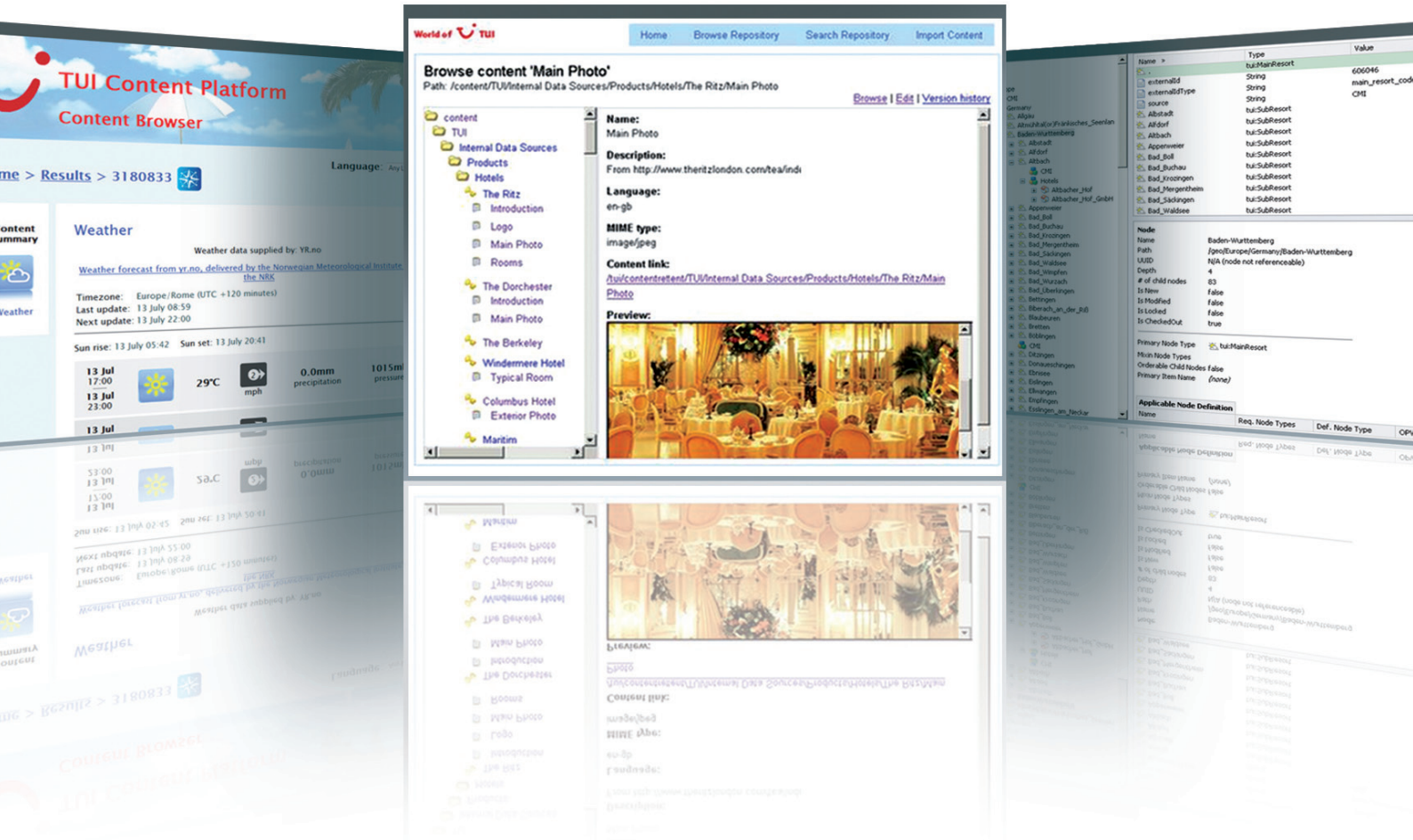
PRİOCEPT

internet technology consultants