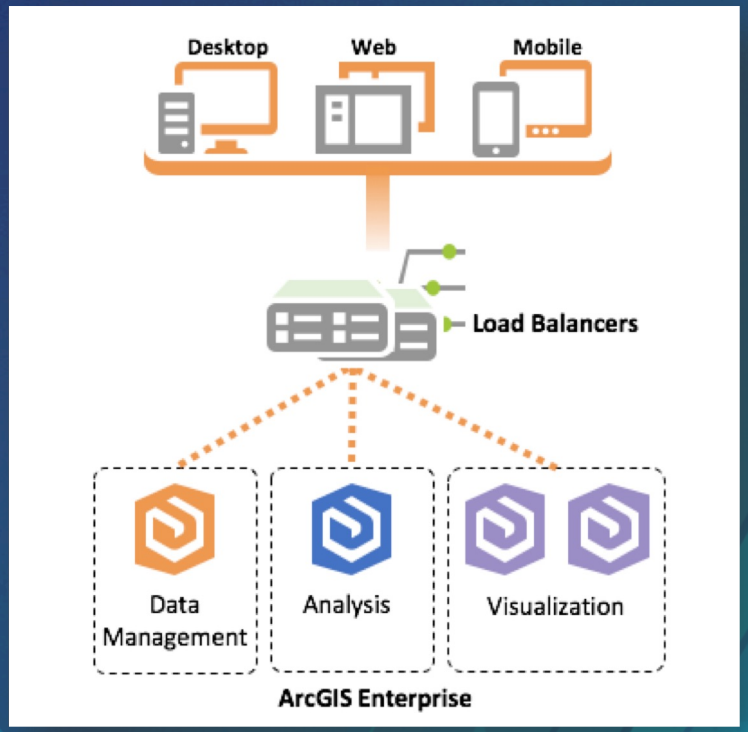
Workload Separation

Antonio Cotroneo & Sarah Scher

Best Practice: Workload Separation

Separate technology deployments by solution pattern

- Improve service delivery
- Reduce risks
- Minimize system resource contention



Typically SLAs determine which server deployments need to be separated.

Data Management

- Consistent Daily Use
- Read/Write and Reconciliation Intensive
- Editing, Versioning, and Archiving

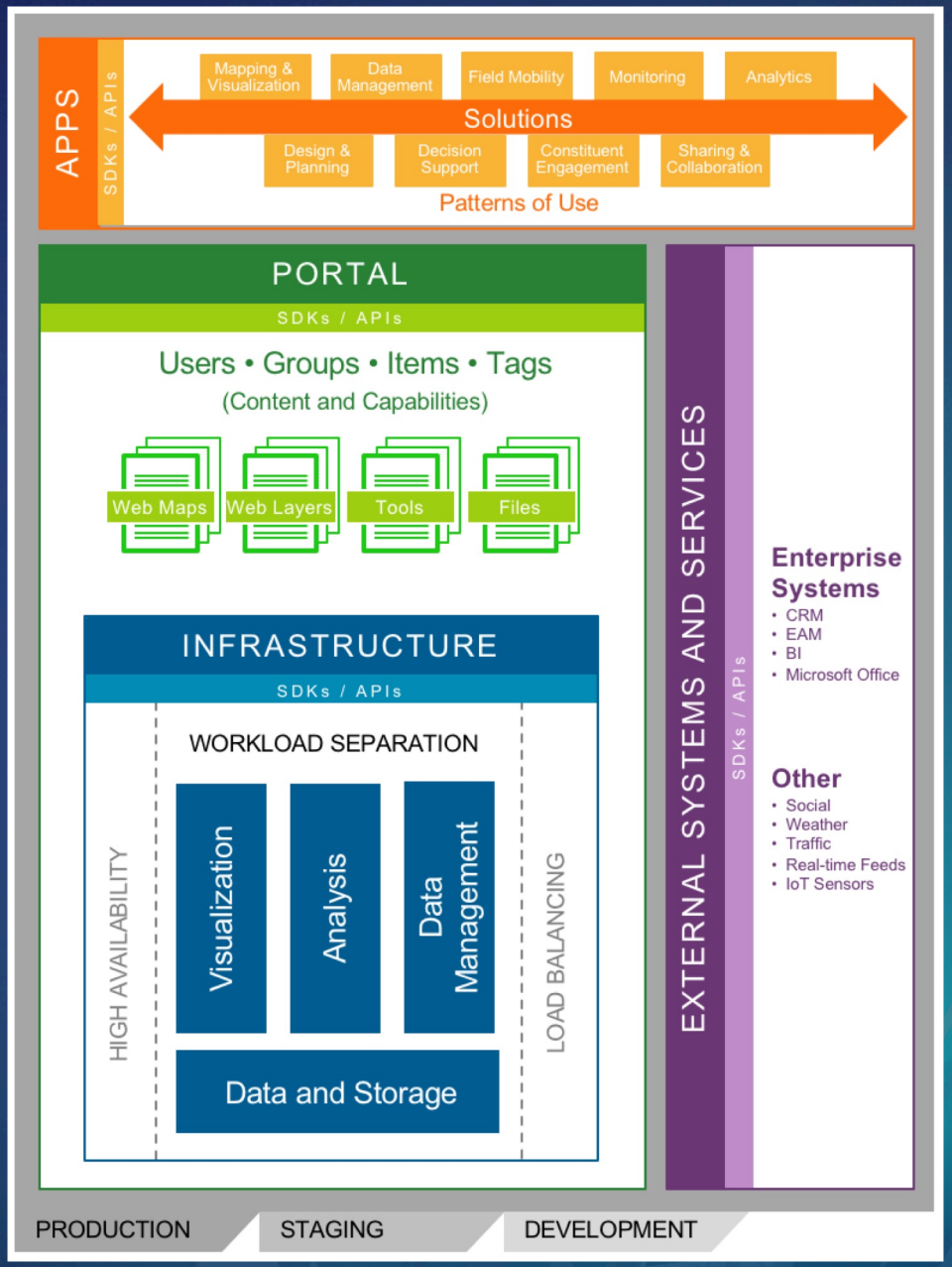
Analysis

- CPU Intensive
- Sporadically Executed Tasks
- Maintained by Lower SLAs
- CPU Idle/Spike

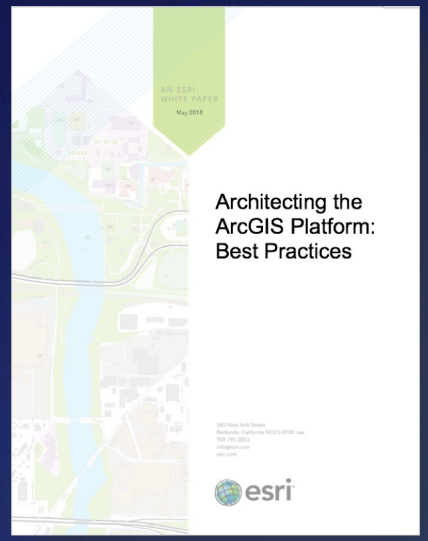
Visualization

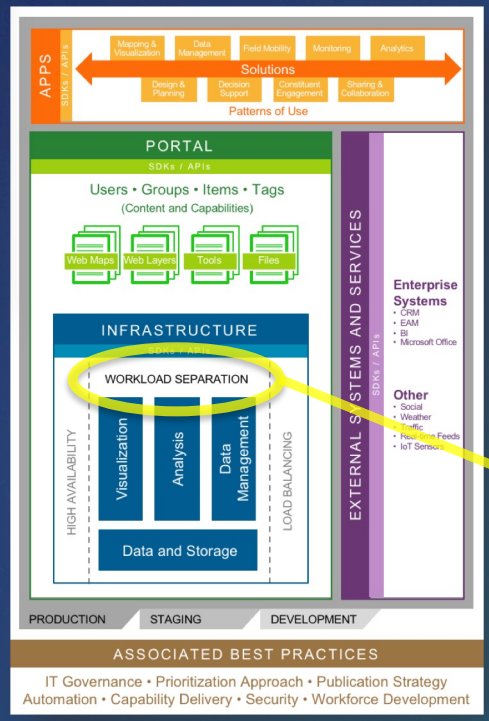
- Less CPU Intensive
- Executed More Consistently
- Maintained By Higher SLAs

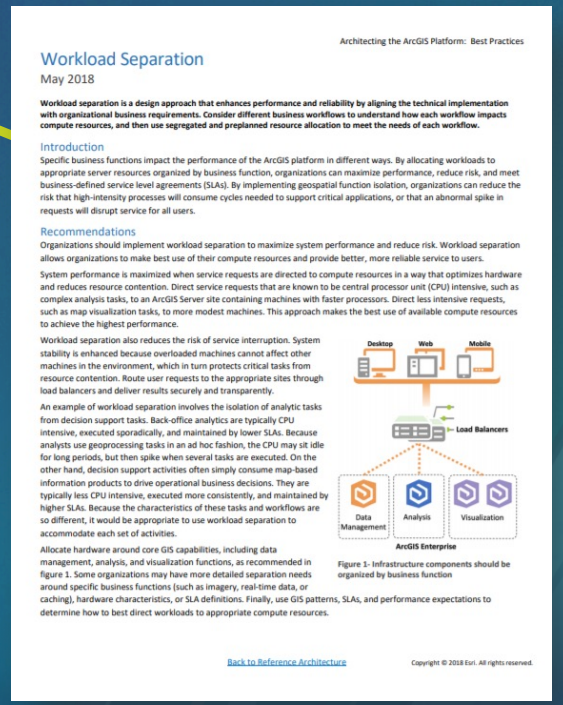
The ArcGIS Conceptual Reference Architecture



The ArcGIS Conceptual Reference Architecture







Architecting the ArcGIS Platform: Best Practices

Workload Separation

May 2018

Workload separation is a design approach that enhances performance and reliability by aligning the technical implementation with organizational business requirements. Consider different business workflows to understand how each workflow impacts compute resources, and then use segregated and preplanned resource allocation to meet the needs of each workflow.

Introduction

Specific business functions impact the performance of the ArcGIS platform in different ways. By allocating workloads to appropriate server resources organized by business function, organizations can maximize performance, reduce risk, and meet business-defined service level agreements (SLAs). By implementing geospatial function isolation, organizations can reduce the risk that high-intensity processes will consume cycles needed to support critical applications, or that an abnormal spike in requests will disrupt service for all users.

Recommendations

Organizations should implement workload separation to maximize system performance and reduce risk. Workload separation allows organizations to make best use of their compute resources and provide better, more reliable service to users.

System performance is maximized when service requests are directed to compute resources in a way that optimizes hardware and reduces resource contention. Direct service requests that are known to be central processor unit (CPU) intensive, such as complex analysis tasks, to an ArcGIS Server site containing machines with faster processors. Direct less intensive requests, such as map visualization tasks, to more modest machines. This approach makes the best use of available compute resources to achieve the highest performance.

Workload separation also reduces the risk of service interruption. System stability is enhanced because overloaded machines cannot affect other machines in the environment, which in turn protects critical tasks from resource contention. Route user requests to the appropriate sites through load balancers and deliver results securely and transparently.

An example of workload separation involves the isolation of analytic tasks from decision support tasks. Back-office analytics are typically CPU intensive, executed sporadically, and maintained by lower SLAs. Because analysts use geoprocessing tasks in an ad hoc fashion, the CPU may sit idle for long periods, but then spike when several tasks are executed. On the other hand, decision support activities often simply consume map-based information products to drive operational business decisions. They are typically less CPU intensive, executed more consistently, and maintained by higher SLAs. Because the characteristics of these tasks and workflows are so different, it would be appropriate to use workload separation to accommodate each set of activities.

Allocate hardware around core GIS capabilities, including data management, analysis, and visualization functions, as recommended in figure 1. Some organizations may have more detailed separation needs around specific business functions (such as imagery, real-time data, or caching), hardware characteristics, or SLA definitions. Finally, use GIS patterns, SLAs, and performance expectations to determine how to best direct workloads to appropriate compute resources.

Figure 1- Infrastructure components should be organized by business function

Back to Reference Architecture

https://go.esri.com/bp


esri
THE
SCIENCE
OF
WHERE