

Technical Perspectives on Scale Up and Scale Out Architectures for Storage Systems

WebTech Q&A Session – May 19, 2010

1. Please explain software inheritance again?

Once a Hitachi or third party OEM storage system is externally virtualized through Universal Volume Manager behind a Hitachi Universal Storage Platform® V or VM all the advanced software feature capabilities typically available for internal storage now become available for the externally attached virtualized storage. For instance, Hitachi Dynamic Provisioning, Hitachi Universal Replication, Advanced Cloning and Snapshots all become available capabilities on the externally virtualized systems once the appropriate software licenses have been enabled. So any system that is externally virtualized inherits the native software capabilities of Universal Storage Platform V and Universal Storage Platform VM

2. What is the impact of scale out on path management?


Horizontal scaling can be used to increase throughput and capacity, but it does so at the cost of additional management, increased host and fabric connectivity, and isolation of components. Because each horizontal scale out unit supports its processors and host and fabric connectivity every time a scale out unit is added, the cost of the whole scale out unit and the additional connectivity it requires must increase. Unlike scale up technologies there exists no abstraction layer allowing capacity to be added without forcing an increase in processing power or connectivity. Scale out technologies must expand in a linear fashion so as capacity is added beyond the size of a scale out unit the requirement to purchase additional connectivity and performance becomes mandatory. There is little flexibility for workloads that scale in a non-linear manner such as Tier-3, which typically require immense capacity but need minimal additional host connectivity and processor performance.

3. How do you adjust for performance variations in scale out architecture?

Scale out architectures need to be managed like multiple workload containers. A workload that easily fits in a scale out unit container creates no issue. However workloads that grow in a nonlinear fashion and start to expand beyond the boundaries of a scale out unit now must be manually juggled and balanced to ensure the host application is achieving maximum performance and not over or under utilizing any scale out unit. Horizontal scale out typically works well initially when workloads are small and can be compartmentalized but over time as they grow they tend to become imbalanced and create an additional challenge to manage and support. If you look at scale out units like parking spaces in a ramp garage you can understand some of the challenges. If the parking spaces were all for compact cars and everyone drove compact cars there would be no issues. However, once a large truck shows up it will overlap into an additional space creating both an over and underutilization effect.

4. Other than scale up being a better fit for OLTP and random data type versus scale out architecture, is there a capacity sweet spot for selecting the right architecture.

Scale out architectures can work well for workloads that can be compartmentalized and grown in a linear fashion (They always need an equal balance of ports, processors and capacity as they grow). Scale out doesn't work well for workloads that grow in a nonlinear manner or have varying degrees of



access intensity over a given day. For instance if host ports and processors are being overrun on a given scale out unit there is no way to address the issue other than to migrate some of the workload to another scale out unit in order to balance things. With a scale up architecture the abstraction layer that exists between the front and back end provides ability to grow in a nonlinear fashion mitigating this type of exposure. So in summary, scale out can work but it takes work and time to continually maintain balance across the architecture. An architecture that can scale up, scale out and also provide embedded external virtualization masks these issues and provides capability to increase throughput and performance without need to migrate workloads.