Business white paper

HP Mobility Management

Managing subscriber data in 3G/LTE/Wi-Fi networks

Table of contents

Managing subscriber data in 3G/LTE/Wi-Fi networks

Mobile carriers face complex business and technology challenges as service offerings rapidly expand and need 100% uptime. Success requires excellence in three areas. New HP solutions uniquely deliver all three and solve the subscriber data challenge while delivering industry best uptime.

Executive summary

"... Carriers are facing two big challenges...

***Technical challenges* in integrating new 4G/LTE services with older 2G/3G networks, and...**

***Business challenges* including a relentless fight for subscribers amid a tight margin squeeze..."**

The challenge
Mobile carriers stand at the confluence of two vectors of challenge in the industry: technology and business. The interplay of these vectors has created an inflection point where distinct strategic choices must be made which will likely have significant impact on future success.

The technology challenges are significant and include integrating diverse 2G/3G with emerging 4G and Wi-Fi networks, migrating to LTE and IP Multimedia Subsystem (IMS), and providing a single view of the subscriber across the full services portfolio – all the while maintaining 100% uptime.

The business challenges are no less acute. These include the relentless fight for subscribers, a need to implement the latest services as rapidly as possible, a difficult competitive margin squeeze and the broad paradigm shift that introduces commercial IT into traditionally custom telco-hardened networks.

Key success factors
To effectively navigate this evolving mobile landscape and leverage this inflection point, mobile carriers must align strategies and resources to deliver three capabilities extremely well:

- Mobility convergence
- Optimized Data Access
- Extreme Availability

Together, these capabilities demand a fundamental shift for carriers: expanding beyond *telco-hardened networks* to add *IT-hardened data access*. They must merge two different worlds.

Carriers that make this shift effectively will be able to deliver unparalleled subscriber service while maximizing revenue from both existing and new mobile services. Those who do not will likely find the co-existence of 2G/2.5G/3G/4G networks increasingly fraught with crisis while customer service levels are increasingly difficult to meet. This complexity will grow with increasing subscribers being serviced.

Mobility convergence
The first key success factor, mobility convergence, is the intelligent integration and simplification of the diverse technological underpinnings of mobile services.

Convergence is best understood by looking at network evolution. Home Location Region (HLR) and Home Subscriber Services (HSS) implementations have continually evolved over time, often leaving carriers with a large number of isolated solutions sometimes from different vendors. In addition networks have expanded with IMS and Authentication, Authorization Access (AAA) capabilities. Too easily these advances turn into stand-alone unconnected technology islands that are costly to operate, hard to manage and difficult to scale.

"... All carriers need intelligent integration...

integrating networks, services, and subscriber data...

the winning carriers integrate by reducing complexity and delivering 100% availability..."

To simplify this technology sprawl, carriers must not only consolidate, but *intelligently integrate* rooms full of legacy systems into a single, centralized, highly scalable solution. All carriers need this intelligent integration – integrating networks, services and subscriber data. The winning carriers integrate by reducing complexity and delivering 100% availability.

These convergence solutions reduce cost, speed time-to-market for new services and help avoid outages that are so quick to grab business-impacting headlines. There are two broad design methods for mobility convergence:

- Maintain & bolt-on
- Integrate & optimize

The maintain & bolt-on approach relies on shallow-integration. It leaves existing networks and solutions predominantly untouched in a "don't fix it if it isn't broken" manner. This approach *adds synchronization complexity* on top of already complex HLR/HSS processing. This approach can succeed in maintaining the status quo - current mobile service stability – but does so at a tremendous cost: it exponentially compounds the pain of growth. It impedes the ability to quickly integrate new services and it gets more difficult as your networks expand. It also compromises service availability by relying on commercial data processing techniques for synchronization.

The Integrate & optimize approach uses deep-integration. This approach *decreases network complexity* by reformulating HLR/HSS services from the ground up. It streamlines core processing in HLR and HSS systems and allows them to share a consolidated subscriber database. Through this simplification, network overhead decreases and choke points are removed. At the same time, failure points found in commercial IT implementations are avoided since processing is housed on fault-tolerant 100% available infrastructure.

The convergence choice is stark. The maintain & bolt-on approach adds complexity in an attempt to avoid changing existing complexity. The integrate & optimize approach decreases complexity through intelligent redesign and by offering simple mix-and-match deployment building blocks - mated pairs, clusters, front-end (FE)/back-end (BE), and N+1. HP offers the only integrate & optimize approach in the market today with such deep integration of app and profile data.

Optimized data access

The second key success factor is data access. HLR/HSS systems hold the all-important subscriber data that allows customers to access mobile services. Unifying this data is vital since it allows carriers to have a unified view of the subscriber and enables a better user experience when subscribers change to a new device, or when they transition between network access types. The Holy Grail is creating a unified subscriber view across diverse networks for a seamless subscriber mobile experience and a database to intelligently up-sell value-added services.

There are two broad design choices for data access:

- Commodity storage design
- Optimized storage design

"... The Holy Grail is creating a unified subscriber view across diverse networks

...facilitating a seamless subscriber mobile experience..."

Commodity storage designs leverage general-purpose IT hardware and software. These flexible components allow solution vendors to more easily create mobile solutions. But this comes at a huge cost: a significant compromise in the speed of data transfer within those solutions. This is often not an issue in commercial solutions but is a significant problem in real-time mobile applications. By their nature, general purpose IT solutions have a long code execution path to satisfy generic commercial processing requirements. For mobile solutions, this long code path translates into an almost intolerable decrease in performance (50X greater latency as compared to optimized approaches) and an increase in unintended choke points that impact throughput.

Optimized storage designs rely on a simple data architecture freed from the weight and complexity of commodity designs. They are lean and simple end-to-end. The HP implementation is an optimized storage design in which HP owns all components of the data path including OS, middleware, database and solution interconnected (with HP ServerNet massive throughput technology). This gives HP the ability to continually offer the most optimized data access possible.

A carrier's choice of data access design is key. Commodity storage designs are weighted with unnecessary functionality and lengthy code-execution paths that can greatly extend the call-path. Optimized storage designs trim all unnecessary processing to delivers the fastest mobile data access available.

Extreme availability

Extreme availability is the third key success factor, and likely the most important. A carrier's worst nightmare is a large network outage that creates a painful mix of grief: negative subscriber churn, painful negative press, and significant out of pocket recovery expenses. But it gets worse - outages can impact quarterly corporate earnings[1], require detailed explanations in annual reports[2], and damage careers[3]. Alarmingly, outages are occurring more often as carriers integrate new architectures based for core network elements based on less-reliable commercial IT into their telco-hardened networks.

There are two broad choices in availability design.

- High availability
- Fault-tolerant

"… Outages: a painful mix of carrier grief…

1-subscriber loss
2-painful negative press
3-recovery expenses
4-annual report visibility

all four can impact corporate earnings and careers…"

High availability designs rely on commercial IT components that have good, but not 100% availability. A detailed analysis of each processing layer (system, data, node and app) reveals failure points in each. As processing is cascaded through these layers, each with availability flaws, the overall solution is increasingly subject to failure. To underscore this fact, after a significant outage, Alcatel-Lucent CEO Ben Verwaayen stated, "Every layer of technology has some issues and lets deal with them"[4]. As one clear example, Linux (found in many high availability designs) has improving availability, but does not reach the 100% level demanded by most carriers.

Fault-tolerant designs create availability from the ground up by ensuring that each component and layer avoids all single points of failure. The solution elegantly recovers if anything fails: a CPU, server, network line, power source, application queue, database record and more. Fault-tolerant designs provide an intriguing paradox: they offer a 100% availability design through patented technology, but deliver it in fully open standards (e.g. SQL, communications protocols, etc.).

To understand this vividly, look for online articles[5] on mobile service outages. You will see a litany of outages with solutions based on high availability designs. These outages occurred during normal processing times. They aren't caused by any natural crisis; they just seem to fail periodically. Compare this to HP fault-tolerant designs that are able to survive earthquakes, tornados, floods and other natural disasters[6] – as well as non-crisis situations.

At $15M/hour cost of downtime, the price of an entire HP solution can be paid for simply through outage avoidance. High availability designs simply cannot bridge the gap to 100% up-time since they included commercial IT components that were never designed for that level of availability.

An integrated HP mobile solution

The new HP Integrated Home Subscriber Servers (I-HSS) provides the best alternative in each of the key success areas: mobility convergence, optimized data access and extreme availability.

I-HSS allows communication service providers (CSPs) to streamline the transition to LTE networks in heterogeneous network (HetNet) environments, embracing a diversity of mobile access networks with a single subscriber management solution. It also helps CSPs implement their unique mobility and subscriber data vision with a set of compelling benefits:

- Best availability –100% uptime design, the only solution with fault-tolerant architecture
- Best TCO – greatly limiting operations, management and acquisition costs
- Most flexible and assured network evolution:
  - Proof: 400M (and counting) subscribers smoothly migrated
  - Flexible deployment building blocks (mated pair, clusters, BE/FE, N+1)
  - Flexible functional support (2G/3G, LTE, IMS, AAA)
- Most flexible UDR deployment
  - Flexible design options: co-located, centralized, distributed
  - Flexible 3rd party app usage (through standard access methods)
  - Flexible data mining platform for subscriber data and services up-sell

1, 5 See Figure 7 – a sampling of service outages
2 Telenor (Norway), 2011 Annual report
3, 4 New Zealand Herald, Feb 24, 2010.
6 See Figure 8 – 100% availability despite crisis

"... The most compelling fact of all...

No HP HLR/HSS solution has ever been replaced,

while HP has replaced numerous competitive HLR/HSS solutions..."

But most compelling of all: *no HP HLR/HSS solution has ever been replaced, while numerous solutions have been replaced with HP mobility solutions.* Why are HP customers so loyal? It's simple: carriers want solutions that don't go down. They want solutions that easily integrate new technologies into their legacy networks. They also want a unified view of subscribers so that they can offer them the best services to drive customer satisfaction and profit. HP offers that today with I-HSS – ready to support your unique mobile vision.

Mobility at an inflection point

Mobility background and evolution

The Home Location Register (HLR) has long been a core element in carrier networks since it is the central repository for all transient and static subscriber information. In the past, HLRs were introduced as part of the network's switching technology. Later HLRs have moved off-switch to provide a separation between these two functions. In many cases, the HLR application still resides on proprietary switch hardware limited in capacity requiring a number of HLRs to provide adequate service for all subscribers.

As new mobile services evolved, so have mobile networks. One such evolution is the transition to an IP-based infrastructure, deploying an IP Multimedia Subsystem (IMS). IMS addresses the challenges of service interdependence, where simple service building blocks are combined to deliver more sophisticated end-user services that integrate multimedia, data, and voice within a single user session.

Mobile networks continue to change with the introduction of Long Term Evolution (LTE) where carriers face the inclusion of yet another network to manage. LTE provides operators a new kind of opportunity—allowing a true all IP connectivity for their clients. Subscriber data can no longer be considered an in-network repository. Subscribers are connecting via different access methods and multiple devices. Thus, operators need to control their service delivery in real-time regardless of the network involved. This architecture calls for HSS implementation that is off-network, and network vendor independent.

This evolution of service and technology has created a number of challenges for carriers.

Market challenges for carriers

Carriers face a mobile market at a crossroads created by a mix of technical and business realities. Technical realities present exploding complexity and escalating requirements including:

- 3G/4G/Wi-Fi/cable integration
- Migration to LTE
- Demand for 100% up-time in traffic storms and natural disasters
- Need to simplify exploding network complexity
- Need for open networks to avoid vendor lock-in
- Need for securely enabling new IP-based services

In the same way, business realities are becoming more complex and demanding:

- Carrier paradigm shift: movement from all custom networks to custom + commodity IT
- People: must add commodity IT expertise/capabilities
- Competitive: intense fight for subscribers
- Financial: margin growth in tough times
- Public Relations: outages making big headlines

Most carriers have determined that to best compete in this complex technical and business environment, they must be successful in three areas:

- Mobility convergence (HLR + HSS + AAA + IMS)
- Optimized data access
- Extreme availability

Mobility convergence

Mobility convergence is the effective integration and co-existence of disparate elements in a mobile network. This includes integrating 3G (HLR), 4G (HSS), and IMS & AAA services to extend reach and ensure consistent mobile service. Convergence is the first strategic imperative mobile carriers must overcome to provide ever-increasing services with assured up-time.

Why mobility convergence matters

Convergence is a strategic imperative for carriers because it is directly tied to carrier profitability. Successful carriers seek to maximize the exploding 4G/LTE revenue stream and while maintaining and protecting the more legacy 2G/3G revenue. Both streams of revenue require different physical networks, so finding ways to merge them into one seamless set of services is vital.

Additionally, convergence is becoming mandatory for competitive viability. At last count 381 operators are investing in LTE technology in 114 countries to satisfy the strong demand for mobile services that only LTE networks can provide, Carriers must either innovate and converge their networks, or become increasingly competitively handicapped.

Lastly, convergence decisions affects many vital portions of a mobile carrier's business including: overall network costs, frequency/severity of outages, agility of adding new services , subscriber confidence/retention and increasingly, the bottom-line quarterly corporate earnings.

"...Mobility convergence is directly tied to carrier profitability...

...it allows carriers to maximize exploding 4G/LTE revenue while protecting existing 2G/3G revenue..."

Convergence overview

Convergence is not easy. The problem that first surfaces in attempts to integrate is the fact that it is difficult for heterogeneous networks to interoperate. Carriers have multiple networks with different network architectures and different network components since each is designed around different mobility services. Intelligence must be created to logically combine disparate functions into a single seamless offering.

In addition to network integration, there is data and application processing integration that must occur. This IT integration typically relies on commercial IT infrastructure, often Linux based, to reduce cost. Although this IT infrastructure is flexible and low-cost, it does not have the stringent up-time requirements of telco-hardened infrastructure. The reliability of the call path is only as good as the weakest link, so the mismatch of availability designs leaves carriers often with less reliable implementations.

Backwards compatibility is also a requirement that must be met during convergence. Even though new services must be implemented rapidly, old services delivered in the same way to the subscriber must be maintained.

A large challenge in convergence is at the database level. Pre-convergence, information about an individual subscriber may be spread across multiple networks without a single overarching view of the subscriber. The lack of an integrated subscriber view makes the delivery of subscriber services in a seamless manner, difficult.

Mobile solution providers offer two broad approaches for convergence: a maintain & bolt-on design, and an integrate & optimize design.

Maintain & bolt-on design

The most common approach to mobility convergence and integration is a maintain and bolt-on design. This design keeps existing HLR & HSS implementation separate with a call path mentality of "don't fix it if it isn't broken". It therefore puts an emphasis on *minimizing change to existing infrastructure* and providing a shallow level of integration between networks and services. Figure 1 is an example of a typical maintain & bolt-on design.

The maintain & bolt-on design relies on adding synchronization complexity on top of already complex HLR/HSS processing – attempting to coordinate services across unpredictable networks with highly distributed data. This approach can often succeed in maintaining current mobile service stability but does so at a significant cost: it exponentially compounds the difficulty of adding new services. Every new service must be analyzed not only in its own right, but also in

terms of the impact of force-fitting that functionality to co-exist with aging infrastructure that was never designed to accommodate it.

This became painfully clear in 2012 as a major US carrier upgraded to 4G LTE which caused nationwide network problems including the loss of 3G service. "Why you would have this kind of an outage right now is a bit concerning," said John Byrne, an analyst at research firm IDC. "The 3G technology has been commercially deployed for 10 years now in one form or another." Maintain & bolt-on approaches not only make it difficult to add new services, but can bring down existing services in the process of upgrade attempts. This isn't surprising since more parts mean greater complexity and more failure points.

"...The maintain & bolt-on design relies on adding synchronization complexity...

... on top of already complex HLR/HSS processing ..."

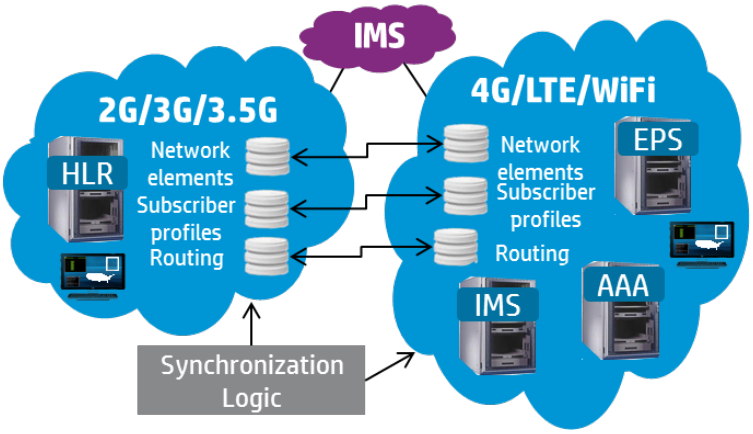

Figure 1 - Maintain & bolt-on convergence design

The maintain & bolt-on approach has numerous shortcomings

- Multiple complex platforms
- Multiple points of failure
- Network element duplication
- Constrained bandwidth, 50X latency
- Data synchronization delays
- Implementation & operations complexity
- Segregated subscriber data in technology silos

Integrate & optimize design

The other major design alternative to achieve integration is integrate & optimize. This design is characterized by a single integrated and streamlined HLR & HSS environment. It rests on the premise of *extensible redesign*: recognizing that integration done well requires extending and modernizing older designs to work harmoniously with newer ones. The HP integrate & optimize approach (See figure 2) offers network flexibility without complexity.

"...The HP integrate & optimize approach offers flexibility without complexity...

...it decreases complexity by reformulating HLR/HSS services from the ground up through deep integration..."

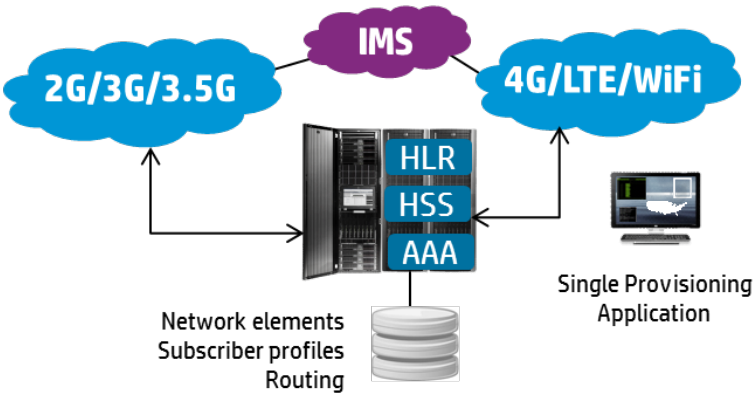

Figure 2 - HP Integrate & optimize convergence design

The HP integrate & optimize convergence design uses deep-integration. This approach *decreases network complexity* by reformulating HLR/HSS services from the ground up. It streamlines core processing in HLR and HSS systems and allows them to easily share a consolidated view of subscriber data on a single server. Through this simplification, network overhead decreases and choke points are eliminated thereby enhancing the subscriber mobile experience. At the same time, failure points – unavoidable with commercial IT implementations– are removed as the entire processing is housed on servers that are designed from the ground up for 100% availability.

This approach offers a number of distinct advantages.

- Simplicity. One platform, one database yields extreme reliability
- Availability. No single point of failure
- Call Speed. Optimized call path, blazing fabric
- Synchronization. One-system immediacy with geo-redundant assurance
- Go-live speed. Fast, single box makes it simple
- Evolution. Fast, changes all in one place
- Unified multi-network subscriber data
- UDC compliance with options to suit diverse customer needs

The best choice for mobility convergence

Both convergence designs add intelligence to harmonize heterogeneous networks. The major difference is that the maintain & bolt-on design puts a premium on maintaining aging networks unaltered whereas the integrate & optimize design puts a premium on modernizing the full integrated network to deliver the best overall subscriber experience. Also, the maintain & bolt-on design *adds integration complexity* on top of existing network complexity. The integrate & optimize design *reduces overall network complexity* by simplifying the heterogeneous network from the inside out.

If a carrier's network is rather static and 4G/LTE growth is not anticipated, the maintain & bolt-on design can be serviceable. However, if a carrier's goal is to drive rapid adoption of 4G/LTE services, then an integrate & optimize design should be considered a requirement.

"...To deliver next-generation mobility, we must create a unified ... subscriber profile.

Only by simplifying and consolidating this data can we leverage it across a growing spectrum of personalized mobility services.

The data is the business..."

Optimized data access

Why data access matters

Transforming core network applications and data repositories into a layered architecture with a common subscriber database for all data repositories is an important approach to reduce operational costs as well as improve scalability and availability.

Implementing an integrated subscriber database enables operations and customer care agents to offer improved support leveraging a unified view of subscriber data; it enables a better user experience when subscribers change from one device to the next, and when they use network intelligent oriented converged services. It also facilitates transition between different network access types, from 2G to LTE, Wi-Fi and WiMax.

In addition, this transformation also enables network operators to play a new key role in the market, to expose a unified view of subscriber data in a controlled way and monetize this valuable asset, facilitating more personalized and advanced services. Also, this valuable subscriber information can be used in real-time analytics to offer tailored services based on data in a subscriber's profile that helps carriers encourage uptake of new services and reduce churn.

Data access overview

Mobile data access is real-time data flow within 2G/3G/4G networks to maintain dynamic mobile service. Of primary importance is a single unified view of a subscriber across diverse networks. Only with this unified view can a holistic and intuitive experience be maintained for the subscriber. The need for real-time access presents a variety of challenges:

- Massive update volumes
- Extreme low latency need
- Rapidly evolving data models
- Data synchronization difficulties

- Commodity IT not built for 100% uptime

To deal with these challenges, two major design alternatives have emerged:

- Commodity storage
- Optimized storage

Commodity storage design
Commodity storage designs typically have the following architectural elements:

- Fault-prone master/slave design with commercial parts
- Only supports UD/IP deployment
- Down time to update, degraded performance on large DB
- High overhead multi-level tree

Figure 3 offers a typical commodity storage design.

"...Commodity storage designs deliver low cost, but also deliver...

commodity performance (with near intolerable latency) and...

commodity availability (unsuitable for 100% uptime)..."

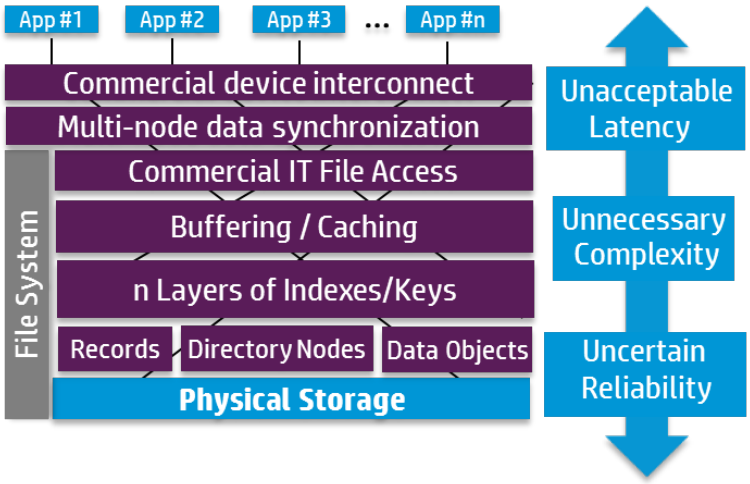

Figure 3 – Commodity storage design

Commodity storage designs leverage a variety general purpose IT components, both hardware and software. Using these components – all proven in commercial IT deployments – solution vendors can more easily create mobile solutions. But this ease-of-solution-creation comes at a huge cost: a significant compromise in the speed of data transfer within those solutions. This is not an issue in commercial solutions but is a significant problem in real-time mobile applications.

By their nature, general purpose IT tools have a longer code execution path to provide the flexibility and breadth of functionality required to give them broad commercial applicability. For mobile solutions, this excess code execution path translates into an almost intolerable decrease in performance (50X greater latency than optimized approaches) and an increase in unintended choke points that impact throughput.

The architecture inherent in commodity storage design presents a number of challenges for mobile carriers:

- Very slow data updates - 50X latency as compared to optimized designs
- Complex redundant dispersed databases
- Compromised data integrity via delayed synch
- Multiple points of data failure & bottlenecks

HP Optimized storage design
Optimized storage designs rely on a simple data architecture freed from the weight and unnecessary complexity of commodity designs. This approach is lean and simple end-to-end and uses the fastest available interconnect technology (HP ServerNet). The HP implementation is an optimized storage design in which HP owns all components of the data path from interconnect

technology, through OS, middleware, database and solution framework. This ensures the ability to continually offer the most optimized data access approach and shortest call path in the market.

Optimized storage approaches are characterized by an active-active design with fault-tolerant componentry offering single system speed and assurance. Figure 4 illustrates the HP optimized storage design.

"…The HP optimized storage design uses customized high-performance telco-tuned technology that embraces open standards…

…freed from the weight and unnecessary complexity of commodity designs…"

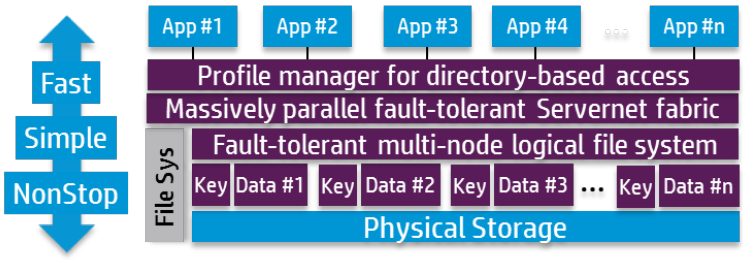

Figure 4 – HP Optimized storage design

Optimized storage designs offer the following benefits:

- Performance. Extremely fast writes with HP ServerNet
- Simplicity. Single consolidated database
- Database Integrity. Real-time sync, fully scalable
- Data Availability. No single point of failure
- Throughput. Maximized by massively parallel design
- 4 interchangeable methods : UD/IP, UD Direct, FE/BE Direct, Dist/Sync

The best choice for data access

Both data access designs seek to unify subscriber data and manage heterogeneous network services. The difference lies in their areas of emphasis. The commodity storage design puts an emphasis on using commercial-grade, inexpensive IT components. Not surprisingly, commodity designs deliver low cost, but also deliver commodity performance (with near intolerable latency) and commodity availability (unsuitable for 100% uptime).

The optimized storage approach rests on a premise that commercial IT technology, no matter how it is packaged, cannot reliably deliver on the 100% availability needs of today's carriers. Optimized storage puts an emphasis on using customized high-performance telco-tuned technology that embraces open standards. Within this context, cost is minimized through streamlined hardware and software. This best-of-both-worlds approach delivers extremely low latency together with simplified and assured data synchronization – all without a cost premium.

Extreme availability

Availability is the bottom line for all mobile networks and the life blood of carriers. Availability is the ability to maintain full end-user mobility service through network failures, traffic storms and even natural disasters like hurricanes, tornados and earthquakes. Availability goes far beyond the ability to recover from failed network components and even full server complexes. It is measured not by the up-time of the componentry, but the up-time and seamless experience of the subscriber.

Why availability matters

Carriers need look no further than the front pages of industry and business press to know the absolute importance of availability. A search for network-outage articles will highlight the searing pain of a carrier's worst outage nightmare – the mix of grief: strong customer dissatisfaction/defection, very visible negative press, and significant out of pocket recovery expenses (estimated at $15M/hour of downtime).

Outages also impact quarterly corporate earnings and careers. Alarmingly, the frequency of outages is increasing as carriers start to integrate less-reliable commercial IT into their telco-hardened networks. Availability is obviously a top concern of carriers today.

Availability overview

There are significant challenges in delivering on the needs of 100% availability in today's mobile networks. These challenges include:

- Skyrocketing downtime cost estimated at > $15M / hour
- Ensuring survivability in network chaos and traffic storms
- Ensuring 100% uptime in multi-vendor commodity IT environments
- Ensuring elasticity during surges
- Delivering subscriber-level (not just hardware) availability
- Intelligent recovery for surges/outages

To deliver 100% availability, two primary designs have emerged. They are:

- High availability design
- Fault-tolerant design

High availability design

High availability designs are the most common design paradigms used by mobility solution providers in their efforts to deliver 100% uptime. Characteristics of these high availability solutions include:

- Leverage availability designs used in commercial applications
- Documented history of failure; many hours to recover
- Documented design flaws (routing table failures, limited by IP network)

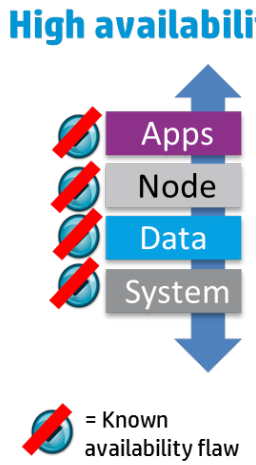

Figure 5 - High availability design

A review of the potential failure points in high availability designs shows that unacceptable design flaws exist at each level of logical processing: at the system, data, node and application levels (see Figure 5). Since availability at any level is no greater than that provided by the layers below it, application level availability – that experienced by the subscriber – for these designs is subject to a troubling interaction between design flaws at multiple levels. Fundamentally, high availability designs fall short of 100% availability because they are not designed to recover from all single points of failure. To underscore this fact, after a significant outage, Alcatel-Lucent CEO Ben Verwaayen stated, "Every layer of technology has some issues and let's deal with them".

It is probably easiest to consider availability from the bottom up. From a system perspective, due to natural hardware component failure rates, disruptions can occur at numerous locations in the server. Although many of these are addressed in high availability designs, many are not. At the OS level, although Linux is maturing in availability, it was not designed for 100% uptime and retrofitting extreme availability on top of commercially-designed software has been problematic.

From a data perspective, high availability designs rely on commodity databases and commodity data infrastructure not designed for 100% availability. The reliability of IP is especially questionable in high availability designs. Routing tables provide unacceptable single points of failure. Also, high availability node designs often implement master/slave configurations for synchronization. Such designs create a dilemma should the master node ever become disabled.

Finally, at the application level, high availability designs often are built for recovery and not for outage avoidance. When built atop system, data and node layers that in themselves are not 100% available, applications cannot deliver full uninterrupted service over time.

Fault-tolerant design

Fault-tolerant design is the less common, but more effective approach to achieve 100% uptime. Fault-tolerant designs have common characteristics:

- Designed for life-critical, real-time needs
- Documented history of best up-time; no single point of failure
- End-to-end failsafe design; not commodity IT, but open IT

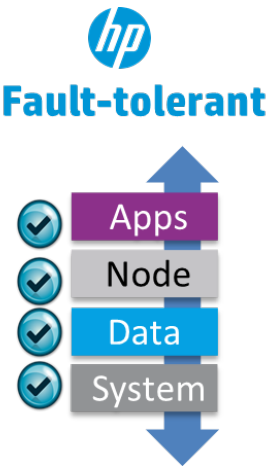

Figure 6 – HP Fault-tolerant design

HP Fault-tolerant designs succeed for the very same reason that high availability designs fail: the fault-tolerance assures that elegant recovery occurs when any single solution component fails. This HP technology has been honed over decades of proven 100% availability in the most demanding processing environments, including market-leading solutions in real-time telecommunications, patient care, banking and capital markets. Figure 6 shows how HP builds fault-tolerance into every level of processing.

The HP fault-tolerant design is based on the premise of distribution of processing across multiple identical components with embedded intelligence that automatically shifts processing from any failed component to one or more like components. This approach allows all components to be productive - actively processing production data continuously – without the overhead of large complexes of unused recovery systems. Systems can be sized to accommodate both peak loads and appropriate processing capacity for failure scenarios.

The HP fault-tolerant design is designed-in at the system level. There are at least two of every hardware component. There are multiple processors and communications paths. Controllers are dual-pathed. The OS is designed so that back-up processors will take over if any copy of the OS fails. Every running process is mirrored in secondary processors should recovery be necessary.

From a data perspective disks are mirrored for both recovery and speed of access, In addition, unique transaction restart capability ensures that data integrity is maintained throughout any potential failure even if processing is shifted between processors and even nodes.

From a node and application perspective the solution environment is designed in a client-server environment that provides continuous service to all solution services.

Today, only HP provides a 100% availability design since only HP provides all levels of processing from the hardware, data base, middleware up through the solution framework. This allows HP to continually ensure the vital interlocked availability design across all levels of processing to deliver full subscriber-level continuity of service. Carriers have flexibility to size their HP processing to provide both the desired steady-state throughput and failure-scenario throughput.

Outages with high availability designs

Performing an on-line search looking for mobile network outages can be informative. Nearly all are damaging to a carrier's bottom-line as well as reputation - providing particularly uncomfortable and unflattering negative visibility. Figure 7 provides a sampling.

Although the list of outages is extensive, figure 7 is only the tip of the iceberg. The table lists outages that carriers are required to reveal due to their magnitude and negative impact. Carriers often experience 5 smaller outages for every 1 reported. As a result, network availability is more tenuous than subscribers typically realize. More importantly, outages represent a significant risk and liability to carriers.

Often these outages are massive – encompassing millions of subscribers. As visibility ramps up on such outages, carriers are caught in a dilemma – do they admit the failure publicly and quickly offer subscriber compensation in order to put the issue to bed, or do they wait to ride out the outage quietly hoping to avoid the visibility that such an admission would create?

In either event the costs are real and significant. When compensation is offered, the out-of-pocket costs are estimated at $15M/hour of outage. When compensation is not offered, still, recovery costs are incurred to bring the network back online. To further compound the issue, carrier reputation and good will is often lessened and a partial loss of the affected subscriber base can ensue.

All of the outages in Figure 7 occurred during normal times (i.e. no driving natural calamity). In addition, all of these outages occurred with solutions that have high availability designs, not the fault tolerant design by HP.

"...Although the list of outages is extensive, it is only the tip of the iceberg...

...carriers often experience 5 smaller outages for every outage they are required to report..."

Availability Model	Impact (# subscribers)	Outage type	When?
High Availability	2.3 million > 24 hr.	Official company apology	Oct '12
High Availability	Millions / 24 hr.	DB outage; £10M cost	July '12
High Availability	26 million /12 hr.	Next-gen HLR outage	July '12
High Availability	3 million / 18 hr. 2 outages	CEO public apology; detail in annual report	June '11
High Availability	Millions / 36 hr.	LTE Network outage	Apr '11
High Availability	Millions/12 hr.	Full network outage due to HLR; impacts profits	Dec '10
High Availability	4 different outages over 3 months	Caused network installer CEO to resign	Feb '10
High Availability	40 million / 4 hr.	Germany & beyond; HLR down, no calls or texts	Apr '09
High Availability	2-4 million / 6 hr.	USA voice/text/data outage	Nov '09

Figure 7 – A sampling of mobile service outages

Up-time with fault-tolerant designs

Figure 8 gives examples of large calamities that HP-mobile solutions survived with no downtime. In each of these cases, significant natural disasters occurred and mobile service was maintained by the HP solution. At these times of natural crisis, mobile service availability can mean more than just carrier reputation – it can mean avoiding the loss of life and property that can result when communication with first responders and those suffering the calamity are disrupted.

"...The HP fault-tolerant design not only avoids down-time in normal times...

... but also avoids downtime in the most severe of circumstances including earthquakes, tornados and floods...."

Availability Model	Disaster	Non-Outage	When?
HP Fault-tolerance		7.4 Mexico Earthquake Continuous HLR operation; CPUs never > 60% busy. See News clip	Mar '12
HP Fault-tolerance		5.8 US Earthquake Tripled CPU usage with NO HLR disruption. See News clip	Aug '11
HP Fault-tolerance		Texas Tornado rips Dallas CPU usage spikes but no HLR outages. See News clip	Apr '12

Figure 8 – 100% availability despite crisis

Considering both figure 7 and figure 8 in combination is interesting. Note that massive failures in high availability solutions have occurred without the presence of a driving natural calamity. These high availability designs should be expected to fall victim to outages even more readily in the event that the network is impacted by natural disaster.

However, the HP fault-tolerant design not only avoids down-time in normal times, but also avoids downtime in these most severe of circumstances.

The best choice for extreme availability
High availability designs are well suited for commercial applications where certain levels of downtime can be rationalized. However, in real-time mobile telecommunications environments, the cost of downtime is simply too large and the business impact too great to risk mobile solutions to high availability designs.

For that reason, HP continues to base its mobile solutions on the proven fault-tolerant 100% availability design that has supported the most important environments in the world: telecommunications, capital markets, emergency 911 systems, patient care applications and more.

HP Mobility Solutions

HP mobility leadership
HP has a decades-long history of mobility solution leadership including:

- 1st off-switch HLR
- 1st multi-protocol (GSM/ANSI) & 1st geo-redundant data sync
- Most flexible deployment & best reporting/management
- Easiest migration from other solutions (over 400M subscribers migrated and counting…)
- Full ANSI compliance (CDMA, TDMA, GSM/ANSI…)
- World's highest HLR/HSS availability,

This leadership is based on an important set of underpinnings that provide unmatched capabilities. These include:

- Network problem detection and correction
- Geo/Logical roaming areas (allows multi-profiling …)
- SIMS/subscriber pre-provisioning (IMSI…)
- Simplified deployment (e.g. USSD Gateway & M2M core…)

Leveraging HP's industry experience and unmatched breadth of products and services HP is offering a truly differentiated and cost-effective solution for the deployment of an HLR/HSS as well as a common subscriber database. This solution includes:

- Leveraging the evolution of HP NonStop platform investment to provide service operators with the lowest cost. This means a simplified, scalable and robust architecture.
- A true multivendor, multi-technology, multi-domain, end to end network solution that will enable service operators to maximize the ROI of the network equipment and retain and attract subscribers through quality of service (QoS) differentiation.
- Bringing all consumers of information in the Common Subscriber Database the data needed quickly, and in an extremely user-friendly way.
- World-class proven program management capabilities that HP has acquired by the addition of EDS to our portfolio

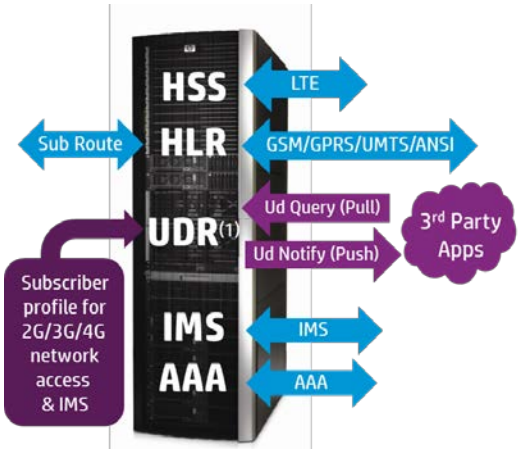

Figure 9 - HP I-HSS solution

HP Integrated Home Subscriber Servers (I-HSS)
The new HP Integrated Home Subscriber Servers (I-HSS) allows communication service providers (CSPs) to streamline the transition to LTE networks in heterogeneous network (HetNet) environments, embracing a diversity of mobile access networks with a single subscriber management solution. As CSPs continue migrating to LTE networks they are tasked with simultaneously managing a wide range of legacy networks accessed by subscribers during the transition. CSPs must be able to easily authenticate subscriber identity and service rights across these heterogeneous networks without interrupting service to subscribers.

I-HSS bridges 2G, 3G, LTE, Wi-Fi and Internet Protocol Multimedia Subsystem (IMS) networks to enable CSPs to centrally manage subscribers' profiles and network access, regardless of the network to which they are connected. By having a single application and real-time synchronized user data repository, CSPs can achieve higher reliability for this mission-critical solution while creating a secured evolution path.

"…I-HSS allows CSPs to streamline the transition to LTE in heterogeneous network (HetNet) environments…

…embracing a diversity of mobile access networks with a single subscriber management solution …"

Uniquely combining the functionality of HP HLR, HP HSS and HP Authentication Authorization Access (AAA), the HP I-HSS is providing standards-based and network-agnostic user data convergence architecture. As a result, CSPs have the freedom to choose their network evolution, to easily enable new services like Voice over LTE and Wi-Fi offload for their subscribers without the need to deploy new network nodes and multiple vendor solutions. Additionally, HP I-HSS enables CSPs to:

- Allow subscribers to move without disruption across network technologies with a single profile, ensuring consistent privileges and service access.
- Lower operation costs associated with concurrently managing legacy and next generation network subscribers by consolidating data access and applications. I-HSS also allows carriers to manage fewer systems with a significant reduction in signaling 8 links
- Ensure high reliability of subscriber data in heterogeneous networks thanks to fault-tolerant architecture and the HP NonStop platform
- Least risky & most flexible deployment solution
  - Carrier-proven fault-tolerant hardware/software stack
  - Side by side evolution with existing network elements for gradual migrations
  - Mix & match best-of-breed design elements
  - Eliminates the new entity introduction risk
- Standards compliant & specific innovation for Subscriber Data Management (SDM) including Traffic Surge Control, VLORA, roam restrict, and optimized messaging

I-HSS leverages HP's industry-leading carrier-grade platforms and a complete suite of feature sets based on 20 years of deployments in 2G/3G and LTE networks — with the expertise to address and tailor the solution to the unique requirements for each customer. I-HSS supports seamless roaming, call delivery and remote feature access on a computer-based platform that delivers linear scalability and unprecedented reliability. It enables subscribers to use intelligent network (IN) services and service features in expanded calling areas from the same platform. It also provides a base upon which new innovative services can easily be deployed.

HP has addressed mobility challenges with a single management environment for all access features and has a leading price/performance ratio based on optimized alignment between software and hardware. The underlying HP NonStop continuously available computer systems are identified as having the lowest total cost of ownership (TCO) in the industry and provide unsurpassed scalability and availability. Scaling to 16 CPU blades, the HP I-HSS can support over 42M subscribers in typical call models on a single node. HP I-HSS nodes are typically deployed as a load shared mated pair or 3- or 4-way configuration for full geographic redundancy and database synchronization, forming a single logical network element.

HP has built up an extensive suite of tools and knowledge regarding the migration of subscribers from incumbent HLR platforms to HP I-HSS. These tools as well as specific knowledge will be assigned to I-HSS implementation projects to help guarantee an effective and seamless subscriber transition. Subscribers have been migrated to HP HLRs from most proprietary HLRs including Alcatel, Nortel, Huawei, Lucent, Siemens and Ericsson HLRs.

Across five continents and in over 35 countries, hundreds of millions of wireless subscribers currently rely on HP technology for their mobile services.

User data convergence (UDC)

Mobile network operators (MNO) continue to expand into multi-technology, multigenerational networks such as with 3G, 4G-LTE, and Wi-Fi, yet they must continue to provide subscribers a consistent service experience regardless of the network to which they are connected or roaming into. The key is the ability to authenticate subscriber identity across networks. Subscriber identity needs to be managed centrally, irrespective of the network the subscriber is connected to at that moment.

HP I-HSS provides a full- featured SDM function that is based on User Data Convergence (UDC) architecture. With a network agnostic architecture, subscriber identity is continuously carried over no matter which access method is first used, allowing delivery of an undisrupted service experience.

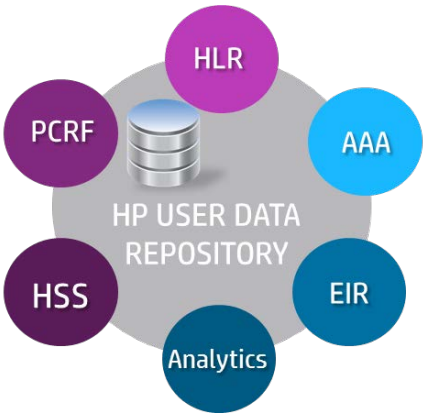

Figure 10 – User Data Repository

As the HP HLR/HSS is a centralized solution from a logical viewpoint, only one subscriber database is implemented. This implies dramatic simplification of routing in the signaling network as all subscribers can be found at the same location. One subscriber database also provides the benefit of having only one provisioning interface to all subscribers, greatly simplifying subscriber management. Furthermore, the HP HLR/HSS is managed as one HLR/HSS entity greatly saving operations cost. Physically, the subscriber database is distributed in at least four different storage devices providing guaranteed data integrity of 100% under all circumstances.

This unified subscriber database is called the HP User Data Repository (UDR). UDR provides a number of benefits:

- NonStop hardware & software platform
  - Unparalleled reliability and fault tolerant
  - Massive scalability benefits
- 3GPP User data (Ud) LDAP interface
  - Open LDAP Server compliant
  - Hierarchical data structures
- HP Optimized Ud interface
  - Available to I-HSS as front-end servers in a NonStop cluster
  - High availability with reduced latency
- 3GPP Ud SOAP Notification
  - Subscription based notification
  - Entry or attribute level subscription
  - Configurable retry persistence

The HP UDR is not restrictive and limited to HLR, HSS and AAA databases. It can also be deployed to host other types of subscriber data for applications like EIR, PCRF, MNP, ENUM and BSF. Currently HP UDR is used to support AAA, MNP and other applications in live operators deployments. See Figure 10.

"...HP UDC - one subscriber database provides the benefit of having only one provisioning interface to all subscribers...

... greatly simplifying subscriber management..."

All back-ends are active at the same time and contain the same copy of the subscriber database. In a distributed architecture, the Application Database Synchronization (ADS) mechanism ensures that the subscriber data in the data nodes, the back-ends, are kept synchronized in real-time. There is no single point of failure and synchronization can start from any back-end. In normal operations all back-ends are active and share the load coming from the front-ends.

ADS sends a copy of the updated data in one back-end to the other back-ends at the time the local database is updated. If a back-end database is unavailable (e.g. due to planned outage of a node for a system upgrade), ADS will queue updates. ADS relies on elaborate client-server mechanisms where one node sends updates to the other nodes through the TCP/IP network. Buffering and queuing mechanisms allow temporary loss of connectivity between the nodes without loss of data.

The HP solution facilitates a smooth transition from 2G/3G to LTE networks and to IMS services. The I-HSS features of HLR and HSS can be active on the same front-end node or can be deployed separately on different front-ends as scale and geography dictate. Their common repository is managed consistently in the back-end UDR and their management environment is integrated as a multi-application and multi-protocol (SS7, Diameter) real-time environment.

Deployment Models

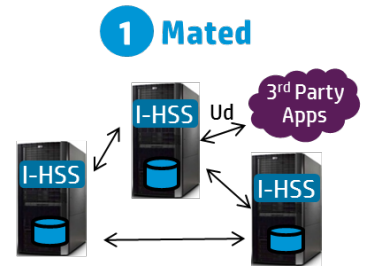

Figure 11.1 Mated servers. Mated Triple shown.

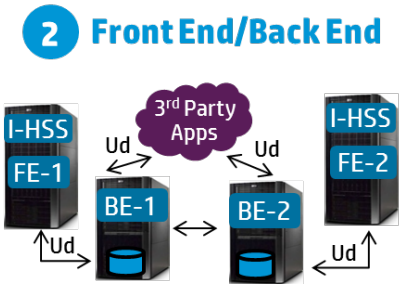

Figure 11.2 Front-End/Back-End Mated pair of FEs/BEs shown.

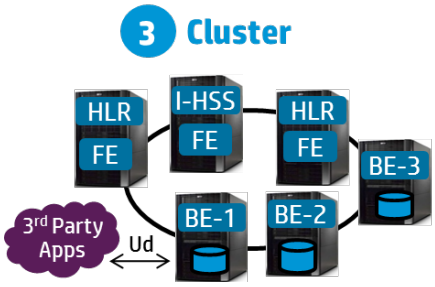

Figure 11.3 Clusters

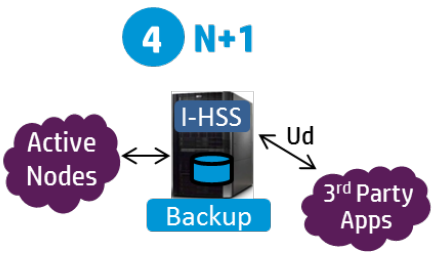

Figure 11.4 N+1

Deployment overview

Carriers require extensive flexibility as they deploy their mobile solutions. The considerations are many: preserving existing network functionality, side by side evolution, sizing for fail-over, minimizing network overhead, synchronizing subscriber data records across geographies, and simplifying network management - just to name a few.

Flexibility is optimized by offering a series of easily-implemented deployment methods - building blocks that can be combined in simple, powerful and almost endless ways. HP offers 4 such deployment methods:

- Mated servers
- Front-end/back-end systems
- Clusters
- N+1

Carriers can select any combination of these deployment methods to create a unique deployment model that best suits their needs. This model can flex and evolve as requirements change.

Mated servers

Mating servers is a method to organize servers for mutual fail-over and synchronization. From a fail-over perspective, servers in a paired environment are sized so that if a failure occurs, the processing of the failed node can be switched to unused capacity in other servers with which it is paired. See Figure 11.1.

Active-active mated servers offer the ability to provide fail-over capacity without the need for stand-by nodes that are dormant until called into action in the event of failure. With HP mobility solutions, servers can be mated in pairs, triples, quads and beyond.

Front-end/back-end

Servers can be organized according to function with certain servers housing subscriber data - back-end servers (BE) - while others support the network and subscriber interconnection - front-end servers (FE).

This separation of functions allows carriers to use different nodes for applications and data, offering more effective management and more precise balancing of processing requirements between functions, especially relevant when the BE data is used by multitude of front-end apps.

Figure 11.2 shows a mated pair FE/BE systems. In this diagram, FE-1 and BE-1 form a logical functional pair, while FE-2 and BE-2 form a separate functional pair. These two functional pairs are mated for recovery. The connectivity options include Ethernet and ServerNet.

Cluster

Clusters allow carriers to create massive complexes of servers within a 65 mile range that have extremely fast interconnect speeds (via ServerNet). Clusters offer massive capacity and flexible logical partitioning amongst servers. Clusters are often used in major cities where mobile traffic is at its highest.

Figure 11.3 shows a cluster of 6 servers all interconnect in a ServerNet ring, a combination of three separate FE/BE pairs.

N+1

N+1 configurations offer stand-alone back-up nodes that can be instantly converted to active processing in the event of a server outage. N+1 designs allow all active nodes to run at maximum capacity with fail-over capacity more broadly managed across a larger set of servers.

Figure 11.4 show an N+1 configuration tied to a network of active nodes.

Beyond the solution

Total cost of ownership

Compelling total cost of ownership (TCO) for mobility solutions is vital. Figure 12 compares the relative cost of various components of mobility TCO between HP and competitors. Although the initial purchase price for HP and competitive solutions are similar, the cost curve skews significantly in HP's favor as the solution is implemented –especially when the dramatic cost of downtime is considered. Solution costs obviously vary by deployment, but through dialog with customers, HP has been able to determine the price bands - High (H), Medium (M) and Low (L) – in which HP and competitive solution TCO components typically reside.

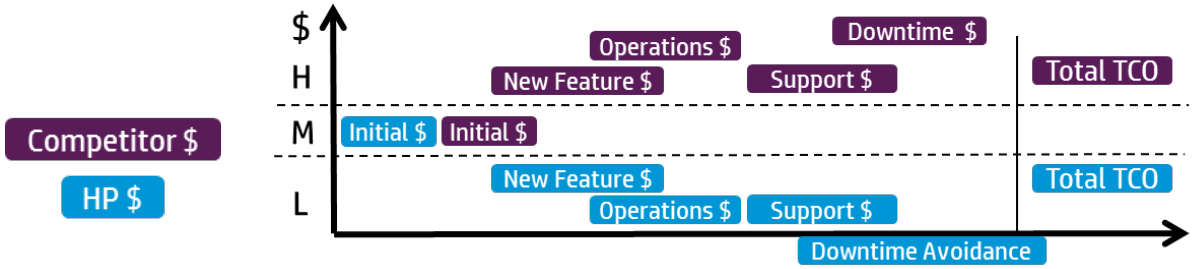

Figure 12 – Mobility TCO[7]

Since HP has full control over its solution stack from hardware deployment to professional services and software features, it has maximum flexibility in implementing new features and can therefore reduce this cost as compared to competitors. In a similar way, HP has full control over support issues and therefore delivers faster resolution which turns into less cost for carriers. Significant automation occurs in HP's fault-tolerant design which means that operations costs for competitive systems are higher, often to a significant degree.

"...With downtime costs of **$15M/hour[8]**

A full HP fault-tolerant mobility solution is often less expensive than the cost of a single outage of a high-availability solution from competitors..."

The most compelling cost however is often the cost of downtime: losing ARPU, offering compensation and eroding brand and market share. This is one area where the solution comparison cost is startling. This is often the highest cost that carriers experience with competitive solutions and the least cost for HP solutions. With 99.9999% and more reliability, a full HP fault-tolerant mobility solution is often less expensive than the cost of a single outage of a high-availability solution from competitors.

Speed of delivery and implementation

HP has an excellent on-time delivery track record in the mobility market - greater than 96% - driven by strict adherence to HP's Standards of Business Conduct (SBC) and Vendor Specific Objective Evidence (VSOE) policies. Some carriers have experienced the "delivery commit game" with other vendors where a late-delivery fee is set and invariably paid – a ploy to simply lock in the sale and include the fee as a cost of doing business.

HP can also implement extremely quickly if required. In one case HP was able to go from contract signing to live LTE HSS implementation in only 10 days.

Customer success

Carriers selecting HP mobile solutions have spoken loudly. Some of the world's largest and most complex networks rely on the HP HLR/HSS:

- More than 35 networks worldwide – GSM/UMTS and CDMA
- More than 400 million active subscribers
- Public references (and many more available per request): Verizon Wireless, Vodacom, H3G, PCL.

As a specialized vendor, HP many times is replacing a solution that was initially deployed as part of an initial network rollout, thus migrating over the years over 150 Million subscribers from many

7 From carrier experienced cost as communicated to HP and available through public sources.

8 IDC "Cost of Downtime Tool", other competitive material and HP internal testing and development over two-year period. © 2009 HP.

different switching environments and competitive HLRs. HP therefore has unmatched expertise in subscriber migrations, an uncommon expertise for most other vendors.

Summary

Carriers stand at an inflection point with their SDM solution implementations, beset by challenges both technical and business. To succeed in this environment, successful carriers require mobile solutions that can deliver on the three key success factors of mobility success:

- Mobility convergence that drives lower operating costs
- Optimized Data Access with open standards
- Extreme Availability

Through intelligent integration for convergence, optimized storage for data access and fault-tolerance for extreme availability, the HP highly specialized offering is best positioned to deliver on the three key success factors for today's mobile landscape. In addition, flexible deployment options, industry leading TCO, best on-time-delivery track record and fast implementations make HP the right choice for carriers who seek a reliable partner to help them fulfill their unique mobile vision.

For more information about HP mobility management go to hp.com/go/HSS.

4AA4-5637ENW, March 2013