Flash Memory SUMMIT

LinkedIn

Impact of NVMe on Servers in Megawebsites

- Server Based Flash Evolves

About the Presenter

Christopher King
Infrastructure Architect (Systems/Storage)
cking@linkedin.com
linkedin.com/in/clhking

Flash Storage

LinkedIn wrote over 475 billion collected metrics to flash yesterday, at a rate of three million writes per second.

Vision

Create economic opportunity for every member of the global workforce

Flash Memory SUMMIT

LinkedIn

The Economic Graph

MEMBERS COMPANIES JOBS SKILLS SCHOOLS KNOWLEDGE

LinkedIn

Mission

Connect the world's professionals to make them more productive and successful

Growing global network

380M+ [1]

Members worldwide

>2 New [1]

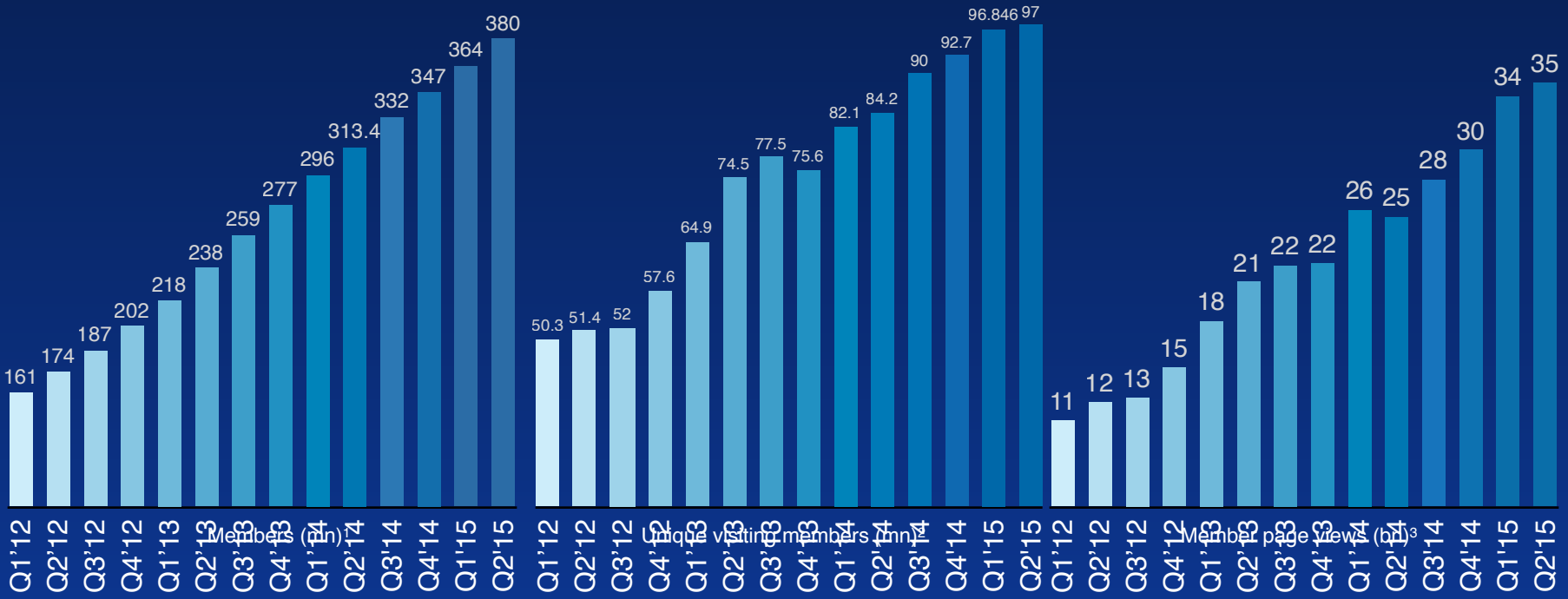
Members per second

35B [2]

unique visiting members

1 As of end of 2Q15 | 2 monthly average during the quarter

Member Value Metrics



1 as of quarter end | 2 monthly average during the quarter | 3 total during the quarter

Flash Deployment

- Flash makes the difference
  - Improve engagement through site speed
  - 300% improvement in commit to disk latency
  - More compact deployment

Flash Stats: PCIe

- Host Based PCIe Flash
  - >1PB deployed and increasing
  - 2x or 3x replication
  - Sizes between 650GB and 4.8TB

Flash Stats: Non PCIe

- SAN/Array based flash
  - <1PB deployed and declining
  - Large RDBMS, Corp-IT Virtualization
- NAS based flash
  - Metadata caching only

Flash Stats (cont)

- Host Based Flash
  - >7500 systems (end of 2010, <50 systems)
  - Average Latency <100µs
- SAN Based Flash
  - <500 systems (end of 2010, <200 systems)
  - Average latency >800µs
- NAS Based Flash
  - <50 systems (end of 2010, 2 systems)

Limitations of PCIe Flash

- Limited set of manufacturers
- Proprietary drivers
- Poor share-ability, if available
- Heat dissipation
- NOT hot swap

NVMe Opens Possibilities

- Drivers
  - Built into all mainstream / updated Operating Systems; that was easy

NVMe Opens Possibilities

- Vendors
  - HGST, Samsung, Seagate, Toshiba/OCZ, Intel, Novachips, SanDisk …

NVMe Opens Possibilities

- Server A has available space; we overestimated.
  - NVMe can be shared
  - We've tested Excelero and it's very promising
  - Mangstor has an interesting offering

NVMe Opens Possibilities

- NVMe comes in 2.5” SFF
  - It gets cooled first
  - Multiple slots
  - The datacenter guys, love it.
    - Current time to move/replace a flash card ~1hr
  - Servers already shipping with 2,10,24 slots
    - 24 slots is a bit silly, but 76TB @ 50µs is cool

NVMe Adoption Planning

- Deploy NVMe in every rack, not per server SKU
- Thin-provision by software or process
- Share between systems of compatible workload
- Mirror data off failing/expiring drives
- Source from as many vendors as can meet performance/reliability testing

Where to from here?

- NVMe now means shareability later
- Test and evaluate the many emerging products
- Narrow down fabric and sharing technology
  - $2 \leq x \leq$ “too many to manage properly”
- Wait for 25Gbit/50Gbit switch silicon to GA

Q/A